```
                                                          Motif 0
human                                        AKFLHWLMSVYVVELLRSFFYVTETTFQKNR
tez1      ISEIEWLVLGKRSNAKMCLSDFEKRKQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNR
EST2      LKDFRWLFISD---IWFTKHNFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVT-
p123      TREISWMQVET-SAKHFYYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFFYVTEQQKSYSK
                                          . *.    .  ....*** **

                                                                Motif 1
human     LFFYRKSVWSKLQSIGIRQHLKRVQLRDVSEAEVRQHREARPALLTSRLRFIPKP--DGL
tez1      TVYFRKDIWKLLCRPFI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKK--NTF
EST2      IVYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLS--NFNHSKMRIIPKKSNNEF
p123      TYYYRKNIWDVIMKMSI-ADLKKETLAEVQEKEVEEWKKS-LGFAPGKLRLIPKK--TTF
            ..*.  *  .    *    *                              .* .**

          Motif 2
human     RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLNYERA
tez1      RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVYMKLLTF
EST2      RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF
p123      RPIMTFNKKIVNSDRKTTKLTTNTKLLNSHLMLKTLKN-RMFKDPFGFAVFNYDDVMKKY
          * *               *                *                    . .

                                  Motif 3 (A)
tez1      KKDLLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS
EST2      KQRLLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN
p123      EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKN
             . *          .*   *   *** . .      .*   *
```

FIG. 1

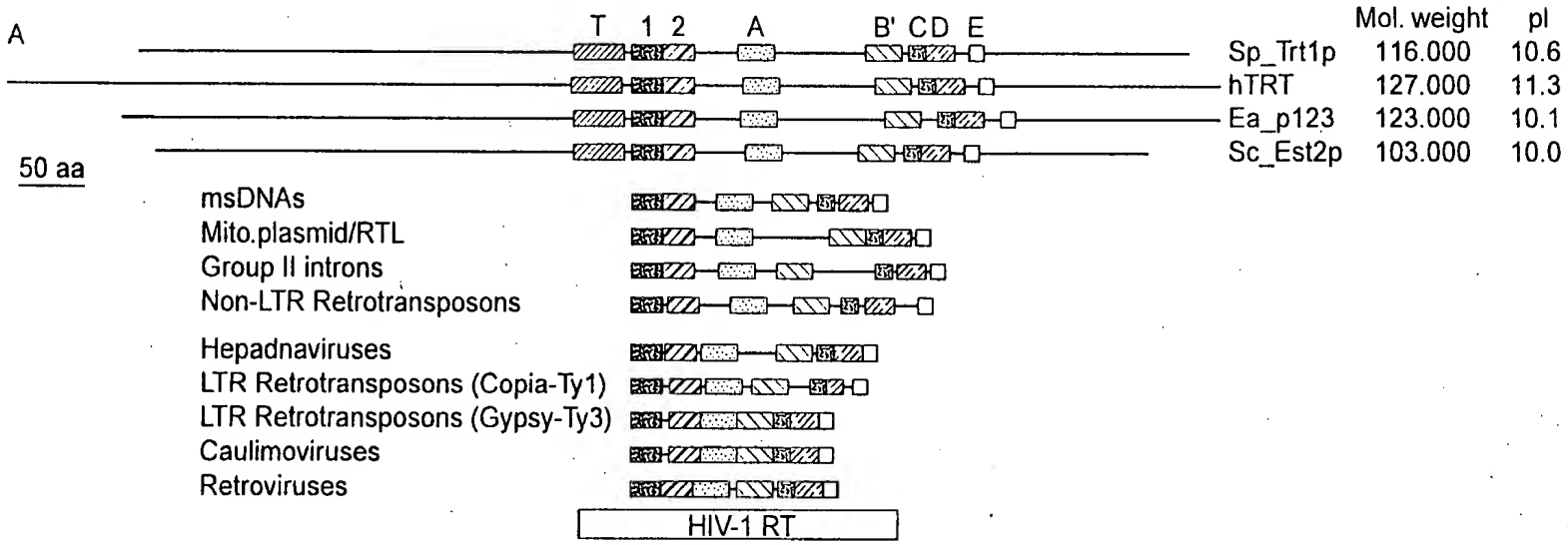


FIG. 2

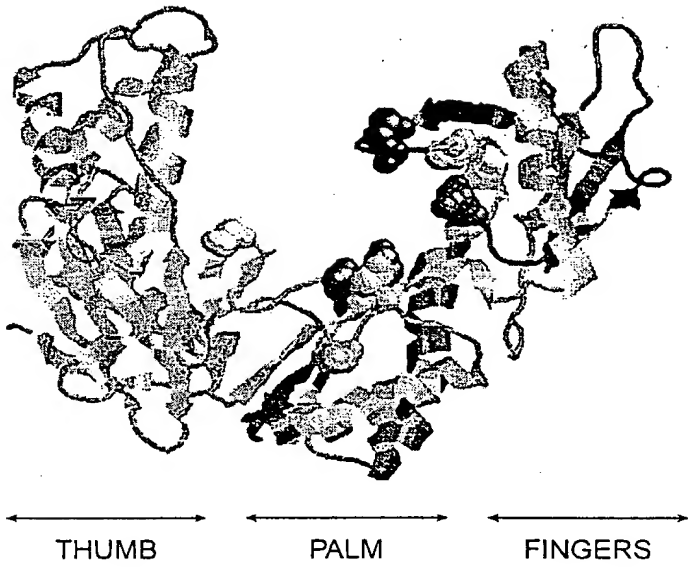


FIG. 3

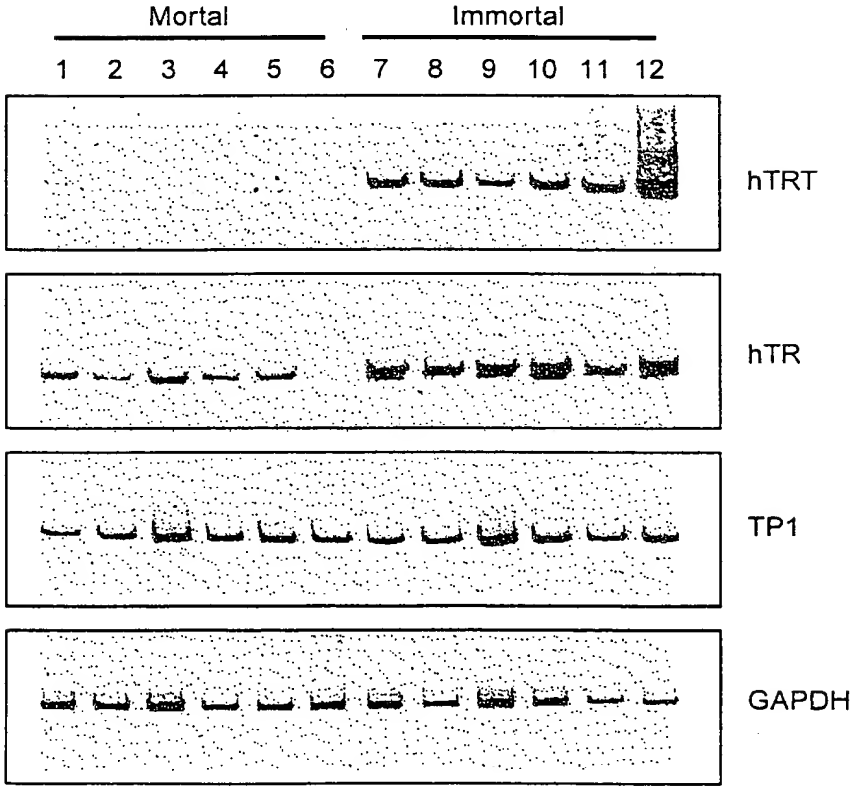


FIG. 5

```
                        Motif T
TRT con         WL     hh hh pFFY TE p    p   Y RK  W  L  h I     K
Sp_Trt1p    429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFITSMKM  8
hTRT        546 WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLK 10
Ea_p123     441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSIADLKK  8
Sc_Est2p    366 WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTWNKLITPFIVEYFK  8

                       Motif 1         Motif 2                                          Motif A
TRT con             h      hRhIPKK p    FRhI  h    h          K               PcLYFh hDh   CYD  I          hhK    K
Sp_Trt1p    NNVRMDTQKTTLPPAVIRLLPKKNT- 0 FRLITNLRKRFLIKMGSNKKMLVSTNQTL 40 FGRKKYFVRIDIKSCYDRIKQDLMFRIVKKKLKD  82
hTRT        EVRQHREARPALLTSRLRFIPKPDG- 0 LRPIVNMDYVVGARTFRREKRAERLTSRV 45 PPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKP  87
Ea_p123     KEVEEWKKSLGFAPGKLRLIPKKTT- 0 FRPIMTFNKKIVNSDRKTTKLTTNTKLLN 41 GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL 100
Sc_Est2p    CRNHNSYTLSNFNHSKMRIIPKKSNN 1 FRIIAIPCRGADEEEFTIYKENHKNAIQP 42 VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN  68
                                                   R                                   AF
RT con                     p hh  h K        hR h       K                       h   hDh   GY  h
Sc_a1       LSNELGTGKFKFKPMRIVNIPKPKGG 0 IRPLSVGNPRDKIVQEVMRMILDTIFDKK 27 FGGSNWFIEVDLKKCFDTISHDLIIKELKRYISD  20
Dm_TART     SILRIGYYPDAWKHAQVKMILKPGKS 6 YRPISLLSGLSKMFERLLLKRLFRVDLFK 32 RKEYCSAVFLDISEAFDRVWHEGLLLKLAKILPY  25
HIV-1       EGKISKIGPENPYNTPVFAIKKKDST 1 WRKLVDFRELNKRTQDFWEVQLGIPHPAG  0 LKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP   7
                               ▲                  ▲

                     Motif B'                          Motif C                      Motif D        Motif E
TRT con     K Y Q  GIPQGS LS hL  h Y DL       F        LLRL  DDFLhIT         A  F    h G  c p   N cK          W G S
Sp_Trt1p    SQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFT  6 LLRVVDDFLFITVNKKD 0 AKKFLNLSLRGFEKHNFSTSLEKTVI 17 KKRMPFFGFSV 181
hTRT        KSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAGI  5 LLRLVDDFLLVTPHLTH 0 AKTFLRTLVRGVPEYGCVVNLRKTVV 19 HGLFPWCGLLL 197
Ea_p123     KFYKQTKGIPQGLCVSSILSSFYYATLEESSLGFL 14 LMRLTDDYLLITTQENN 0 AVLFIEKLINVSRENGFKFNMKKLQT 23 QDYCDWIGISI 179
Sc_Est2p    KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK  8 ILKLADDFLIISTDQQQ 0 VINIKKLAMGGFQKYNAKANRDKILA 20 KELEVWKHSST 146
                                                          F
RT con                hPQG    pP hh    h               h  Y  DDhhh                       Gh h   cK   h         hLG  h
Sc_a1       TYHKPMLGLPQGSLISPILCNIVMTLVDNWLEDYI 55 YVRYADDILIGVLGSKN 2 KMIKRDLNNFLNS-LGLTMNEEKTLI  4 ETPARFLGYNI
Dm_TART     RAGQIGAGVPQGSNLGPILYSIFSSDMPLPHIYHP  7 LSTYADDTIVLSSDILA 6 NENYLKTFSDWADKWGISVNAAKTGH 25 ESKQSYLGVIL
HIV-1       GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFKKQN  4 IYQYMDDLYVGSDLEIG 1 HRTKIEELRQHLLRWGLTTPDKKHQK  0 EPPFLWMGITL
              ▲                                          ▲    ▲
```

FIG. 4

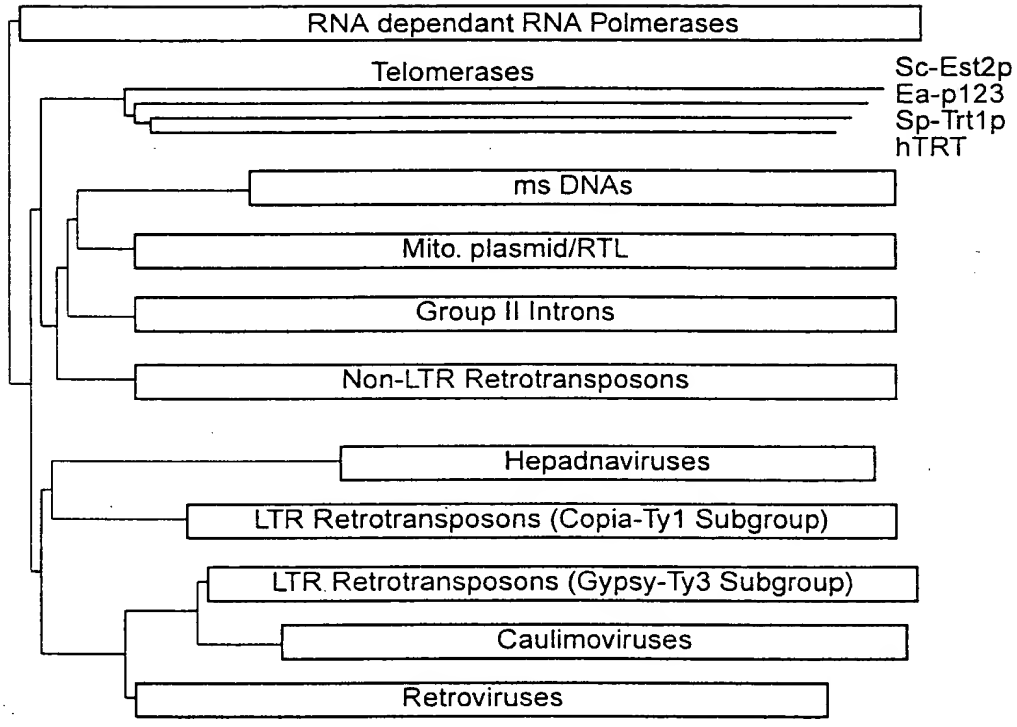


FIG. 6

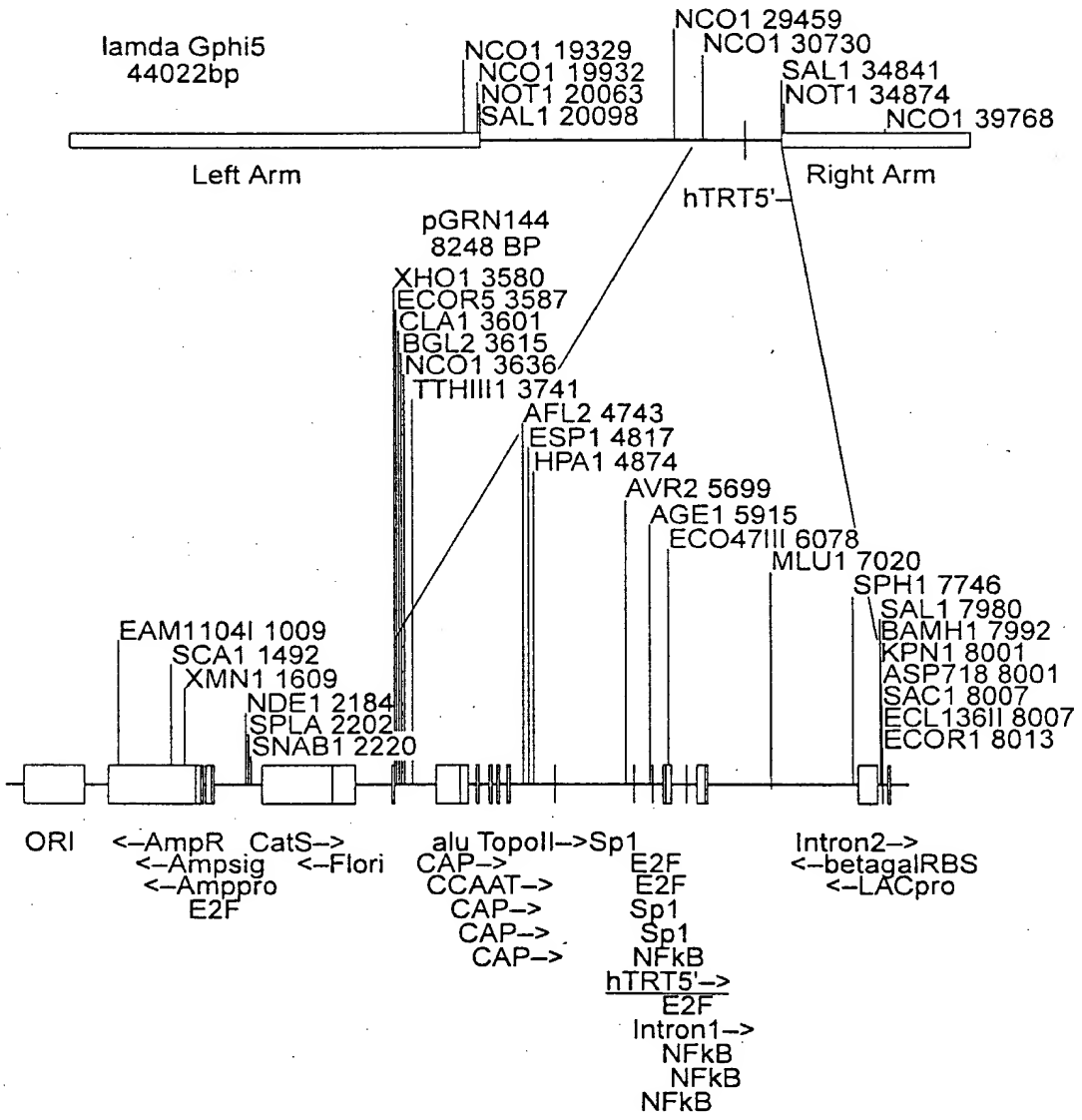


FIG. 7

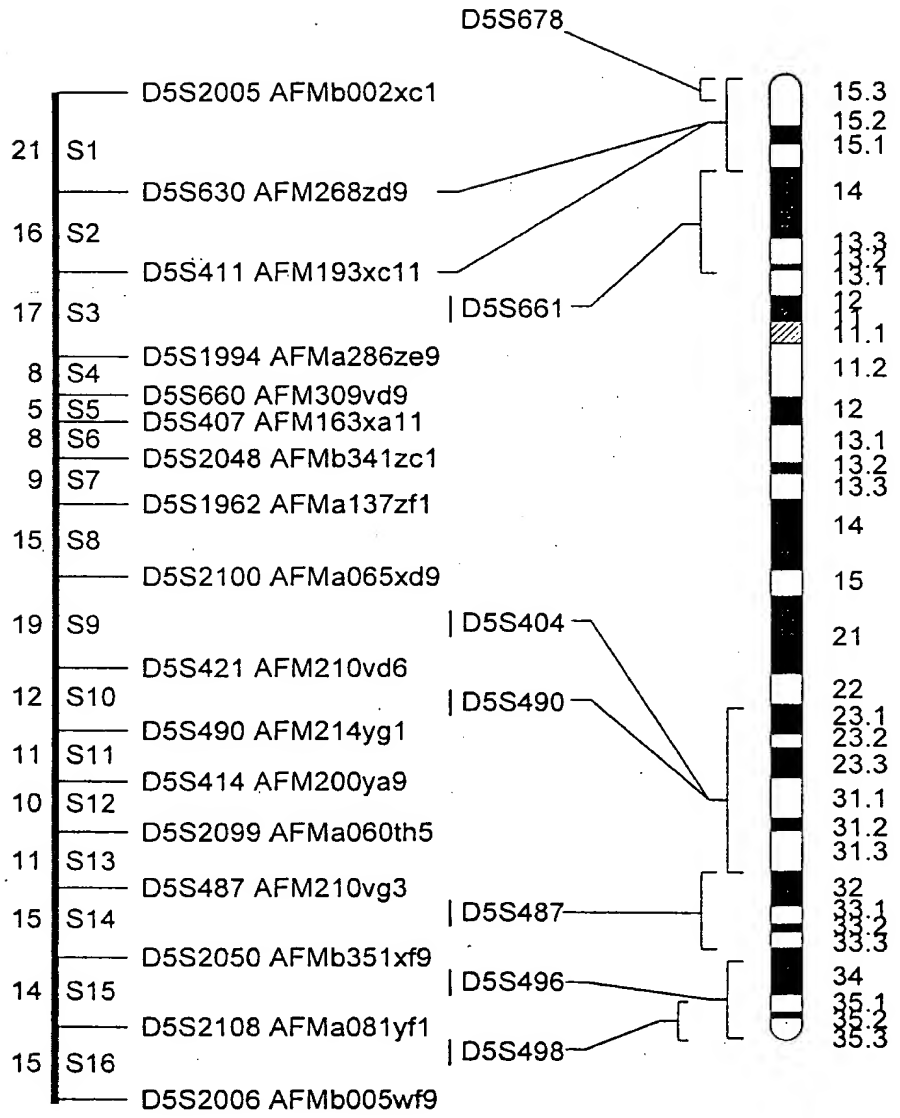


FIG. 8

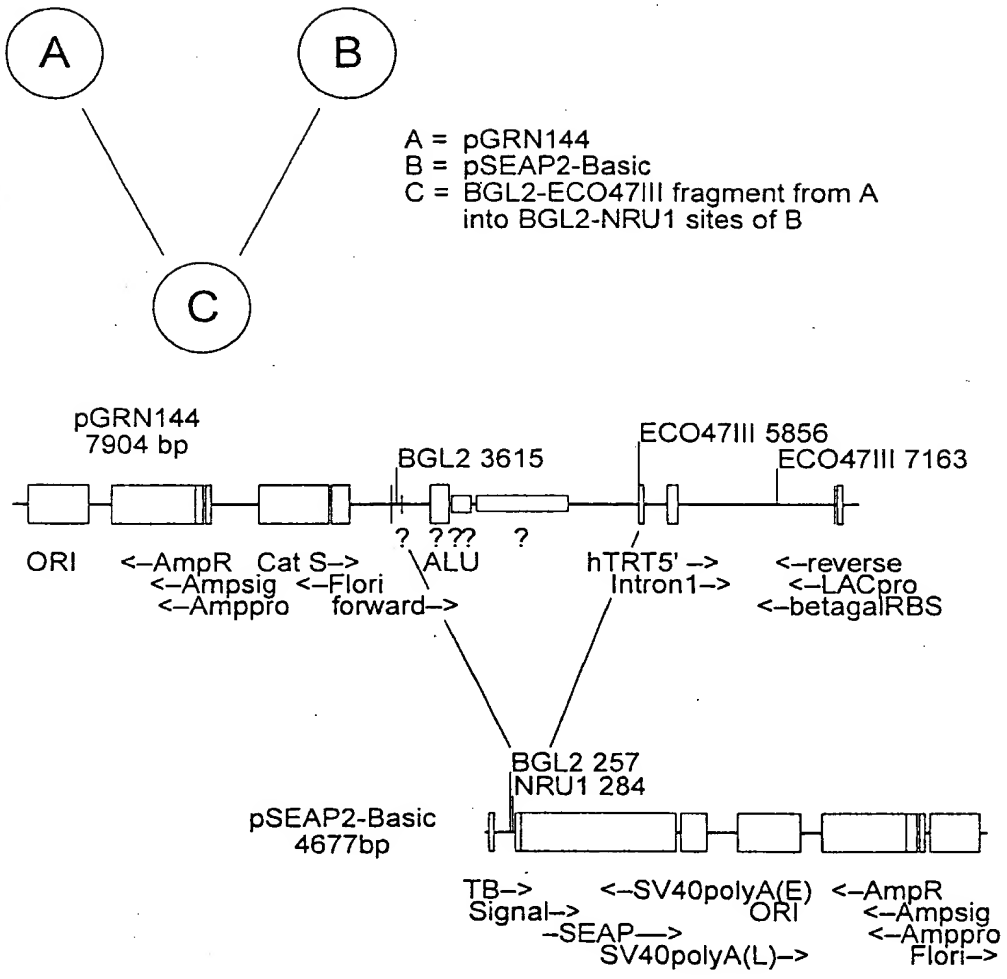


FIG. 9

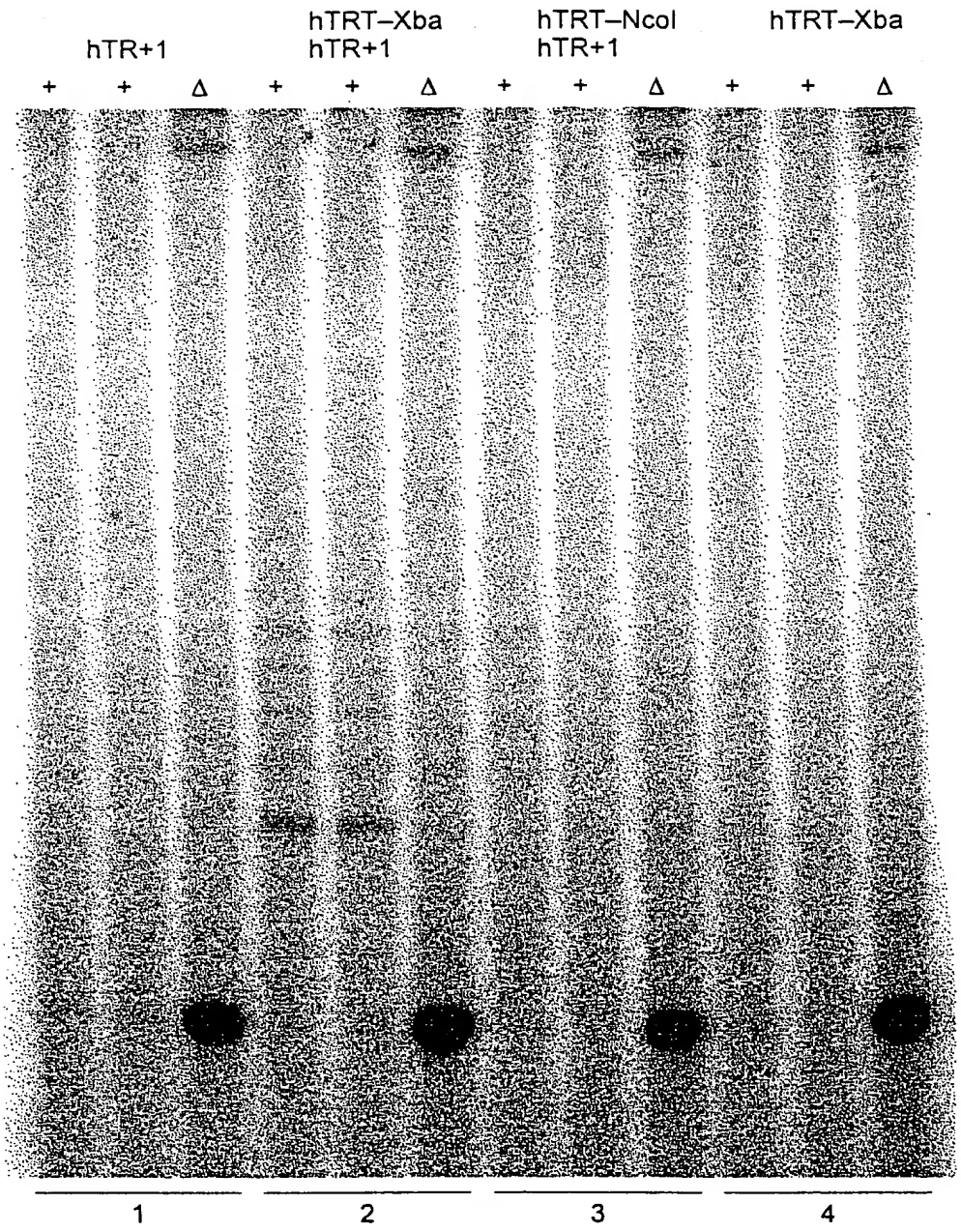


FIG. 10A

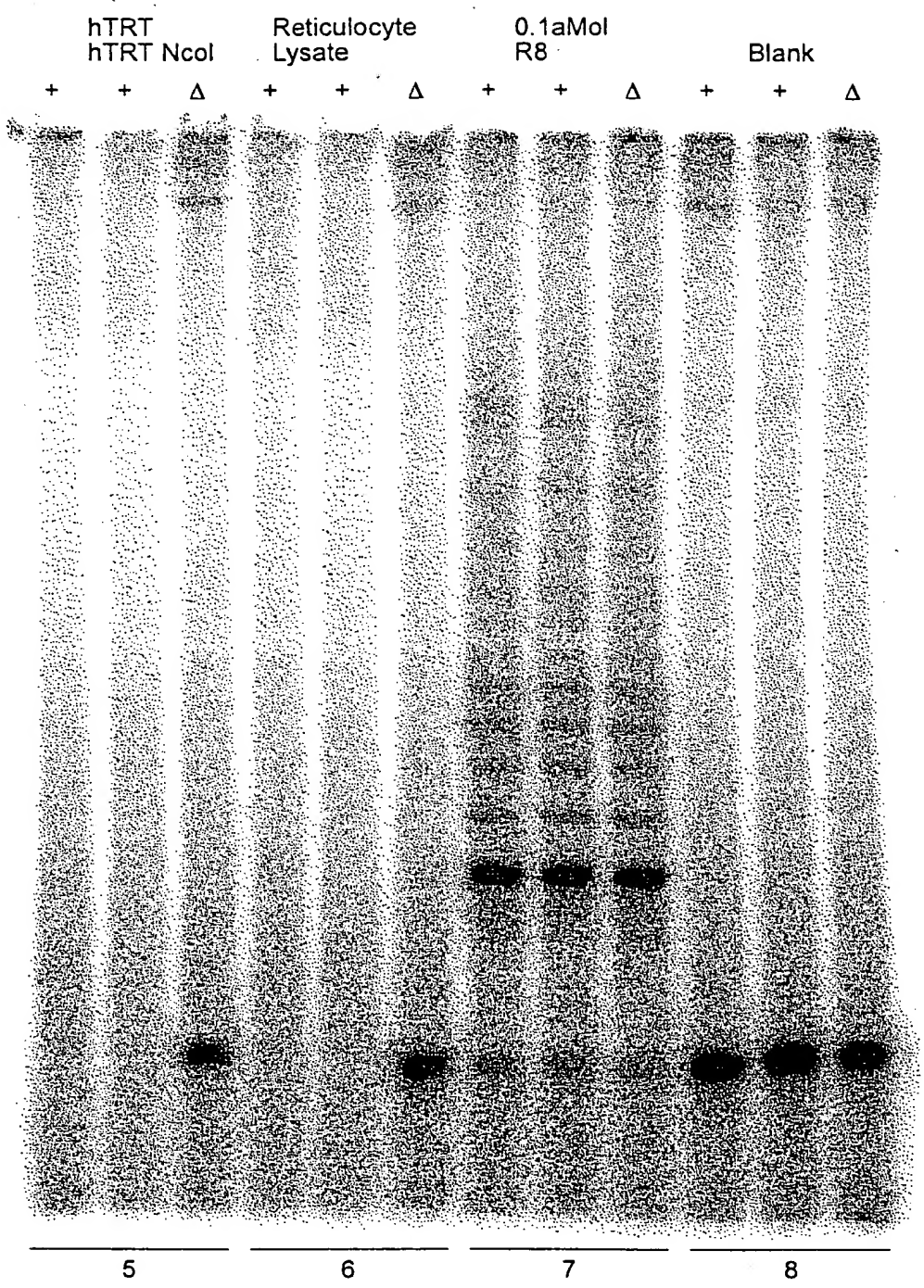


FIG. 10B

Telomerase Specific Motifs

		MOTIF T		MOTIF T'
TRT con		Wl FFY TE y Rk W l I		E V
hTRT	546	WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGI	13	EAEVR
spTRT	429	WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFI	12	ENNVR
Ea_p123	441	WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSI	12	EKEVE
Sc_Est2	366	WLFRQLIPKIIQTFFYCTEISSTVT.IVYFRHDTWNKLITPFI	9	ENNVC

Telomerase RT Motifs (Fingers)

		MOTIF 1		MOTIF 2		MOTIF A		MOTIF B'
TRT con		R iPKk		fR I		p lyF D cYD i		Y q GipQGs lS l y
hTRT	11	SRLRFIPKPDG	0	LRPIV	69	PELYFVKVDVTGAYDTI	104	YVQCQGIPQGSILSTLLCSLCY
spTRT	10	AVIRLLPKKNT	0	FRLIT	66	RKKYFVRIDIKSCYDRI	99	YLQKVGIPQGSILSSFLCHFYM
Ea_p123	10	GKLRLIPKKTT	0	FRPIM	67	PKLFFATMDIEKCYDSV	117	YKQTKGIPQGLCVSSILSSFYY
Sc_Est2	13	SKMRIIPKKSN	2	FRIIA	68	PELYFMKFDVKSCYDSI	85	YIREDGLFQGSSLSAPIVDLVY
RT con		p hh h K		hR h		h hDh AF h GY		hPQG pP hh h

Telomerase RT Motifs (Palm, Primer Grip)

		MOTIF C		MOTIF D		MOTIF E	
TRT con		lllrl DDfL it		g n K		w g s l	
hTRT	15	LLLRLVDDFLLVT	15	GVPEYGCVVNLRKTVV	24	WCGLLLDTRTL	192
spTRT	16	VLLRVVDDFLFIT	15	GFEKHNFSTSLEKTVI	22	FFGFSVNMRSL	176
Ea_p123	24	LLMRLTDDYLLIT	15	VSRENGFKFNMKKLQT	28	WIGISIDMKTL	174
Sc_Est2	18	LILKLADDFLIIS	15	GFQKYNAKANRDKILA	25	WKHSSTMNNFH	141
RT con		h Y DDhhh F		Gh h cK h		hLG h	

FIG. 11

```
181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
    CCTGGGCCGCCGAAAGGCGCGCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCG

                                                NFkB_CS1
                                               GGGRQTYYQC
                                              NFkB-MHC-I.2
                                              TGGGCTTCCCC
                                             ***************************
241 ACGGCCGCCCCCCGCCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCG
    TGCCGGCGGGGGGCGGCGGGGGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGC

                                Intron1
    ************************************************************
301 GCTGGGGTTGAGGGCGGCCGGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
    CGACCCCAACTCCCGCCGGCCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

          NFkB_CS1
         GGGRQTYYQC
          NFkB_CS2
        RGGGRMTYYCC
            Topo_II_cleavage_site
              RNYNNCNNGYNGKTNYNY
    *****************>
361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG
    TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC
```

FIG. 12

```
   1  AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
  51  AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
 101  ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
 151  TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
 201  TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
 251  TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
 301  AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
 351  CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACTTCTT
 401  TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA
 451  ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
 501  CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
 551  ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
 601  GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
 651  ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
 701  GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
 751  GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
 801  ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
 851  AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC
 901  AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
 951  GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
1001  CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
1051  TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAACTTGA
1101  TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
1151  TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
1201  CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
1251  AAAAGAAAGT TAAGAAATAT GTGGAACTAA ACAAGCATGA ACTCATTCAC
1301  AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351  GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGATCAC GAAAACATCT
1401  ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451  CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501  AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551  TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601  GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
1651  ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701  TAAATTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG
1751  AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801  TTTTGGATTC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851  AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTTGCA
1901  ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAACTATC
1951  AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA
2001  CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
2051  TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101  TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
2151  AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201  AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
2251  CCAATATAAT TACATTAACT TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301  GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC ATTTTATTAT
2351  GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA
```

FIG. 13A

```
2401  CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451  TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501  ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551  GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601  GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651  TCAATTGATA TGAAAACTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701  AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751  CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801  CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACTCT
2851  CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901  AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC
2951  TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001  CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
3051  ACTTTTTCCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
3101  AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151  GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201  CTATTCTAAC TTATTTTGGA AAGTTAATTT TCAATTTTTG TCTTATATAC
3251  TGGGGTTTTG GGGTTTTGGG GTTTTGGGG
```

FIG. 13B

```
   1  MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
  51  LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101  SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151  IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
 201  ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTR IFYCTHFNRN
 251  NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAYM
 301  LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351  YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHELIH
 401  KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKLL RWIFEDLVVS
 451  LIRCFFYVTE QQKSYSKTYY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
 501  EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL
 551  NSHLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601  TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651  FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLIVEAKQ
 701  RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY
 751  ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIEKL
 801  INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851  SIDMKTLALM PNINLRIEGI LCTLNLNMQT KKASMWLKKK LKSFLMNNIT
 901  HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
 951  LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
1001  KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A
```

FIG. 14

```
   1 ggtaccgatttactttcctttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaaga  80
  81 actcaataacaataccaagtcaaattccaatatgaaggtgttattagtgatcgataatatttctattttatcggtcgtta 160
 161 ccaagtataaggacaaaaagaacaacttccttcccccctaaagacttttacttttattaatttacttttcaaatatatttcg 240
 241 ggttcgcttacttttaatcgtggtactgttttagctgctacttctagccaaccgcgtgtttctaccccgtcattggatat 320
 321 agctcttggagtagctcacagaaatccttacaaatcttctgatgagactatattagattcattacagtccgtgcatattc 400
 401 ttaacatggagccttacactttagatgagtcacgtcgcatgatggagtatttggtatcatccaacgtttgccttgaaaag 480
 481 gttgataattatttgcaaaatcatgtccttagtggtggtaatccgcgaaagtttttgatgcttgcacacgtctagcatg 560
 561 attgagatattcaaaaatttctatccactacaactcctttaacgcggttttatttttctatttctattctcatgttgtt 640
 641 ccaaatatgtatcatctcgtattaggctttttccgttttactcctggaatcgtacctttttcactattcccccctaatga 720
 721 ataatctaaattagtttcgcttataattgatagtagtagaaagattggtgattctactcgtgtaatgttattagtttaaa 800
 801 gatactttgcaaaacatttattagctatcattatataaaaaaaatcctataattataaatattaatcaatatttgcggtc 880
 881 actatttatttaaaacgttatgatcagtaggacactttgcatatatatagttatgcttaatggttacttgtaacttgc   958

 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA  1018
   1 M   T   E   H   H   T   P   K   S   R   I   L   R   F   L   E   N   Q   Y   V    20

1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG  1078
  21 Y   L   C   T   L   N   D   Y   V   Q   L   V   L   R   G   S   P   A   S   S    40

1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT  1138
  41 Y   S   N   I   C   E   R   L   R   S   D   V   Q   T   S   F   S   I   F   L    60

1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA  1198
  61 H   S   T   V   V   G   F   D   S   K   P   D   E   G   V   Q   F   S   S   P    80

1199 AAA TGC TCA CAG TCA GAG gtatatatatttttgttttgatttttttctattcgggatagctaatatatgggcag 1272
  81 K   C   S   Q   S   E                                                           86

1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA  1332
  87 L   I   A   N   V   V   K   Q   M   F   D   E   S   F   E   R   R   R   N   L   106

1333 CTG ATG AAA GGG TTT TCC ATG gtaaggtattctaattgtgaaatatttacctgcaattactgtttcaaagaga 1405
 107 L   M   K   G   F   S   M                                                       113

1406 ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT  1469
 114                     N   H   E   D   F   R   A   M   H   V   N   G   V   Q   N   128
```

FIG. 15A

```
1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA  1529
 129 D   L   V   S   T   F   P   N   Y   L   I   S   I   L   E   S   K   N   W   Q    148

1530 CTT TTG TTA GAA AT gtaaataccggttaagatgttgcgcactttgaacaagactgacaagtatag T ATC GGC 1601
 149 L   L   L   E   I                                                     I   G    155

1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC  1661
 156 S   D   A   M   H   Y   L   L   S   K   G   S   I   F   E   A   L   P   N   D    175

1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG  1721
 176 N   Y   L   Q   I   S   G   I   P   L   F   K   N   N   V   F   E   E   T   V    195

1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA  1781
 196 S   K   K   R   K   R   T   I   E   T   S   I   T   Q   N   K   S   A   R   K    215

1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT  1841
 216 E   V   S   W   N   S   I   S   I   S   R   F   S   I   F   Y   R   S   S   Y    235

1842 AAG AAG TTT AAG CAA G gtaactaatactgttatccttcataactaattttag AT CTA TAT TTT AAC    1907
 236 K   K   F   K   Q   D                                         L   Y   F   N      245

1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG  1967
 246 L   H   S   I   C   D   R   N   T   V   H   M   W   L   Q   W   I   F   P   R    265

1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA  2027
 266 Q   F   G   L   I   N   A   F   Q   V   K   Q   L   H   K   V   I   P   L   V    285

2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA  2087
 286 S   Q   S   T   V   V   P   K   R   L   L   K   V   Y   P   L   I   E   Q   T    305

2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT  2147
 306 A   K   R   L   H   R   I   S   L   S   K   V   Y   N   H   Y   C   P   Y   I    325

2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG  2207
 326 D   T   H   D   D   E   K   I   L   S   Y   S   L   K   P   N   Q   V   F   A    345

2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA  2267
 346 F   L   R   S   I   L   V   R   V   F   P   K   L   I   W   G   N   Q   R   I    365
```

FIG. 15B

```
2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaatttattaccactaacgattttaccag AC CTC GAA ACT 2336
 366 F   E   I   I   L   K   D                                          L   E   T   375

2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F   L   K   L   S   R   Y   E   S   F   S   L   H   Y   L   M   S   N   I   K   395

2397 gtaatatgccaaatttttttaccattaattaacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
 396                                        I   S   E   I   E   W   L   V   L   G   405

2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K   R   S   N   A   K   M   C   L   S   D   F   E   K   R   K   Q   I   F   A   425

2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E   F   I   Y   W   L   Y   N   S   F   I   I   P   I   L   Q   S   F   F   Y   445

2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I   T   E   S   S   D   L   R   N   R   T   V   Y   F   R   K   D   I   W   K   465

2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L   L   C   R   P   F   I   T   S   M   K   M   E   A   F   E   K   I   N   E   485

2706 gtattttaaagtatttttgcaaaaagctaatattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486                                        N   N   V   R   M   D   T   Q   K   T   495

2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
 496 T   L   P   P   A   V   I   R   L   L   P   K   K   N   T   F   R   L   I   T   515

2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaatttttggtcatcaatgtactttacttctaatctatta 2906
 516 N   L   R   K   R   F   L   I   K                                              524

2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525         M   G   S   N   K   K   M   L   V   S   T   N   Q   T   L   R   P   V   542

2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A   S   I   L   K   H   L   I   N   E   E   S   S   G   I   P   F   N   L   E   562

3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V   Y   M   K   L   L   T   F   K   K   D   L   L   K   H   R   M   F   G       581
```

FIG. 15C

```
3089 tatataatgcgcgattcctcattattaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA   3155
 582                                        R   K   K   Y   F   V   R   I   D   I     591

3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC   3215
 592 K   S   C   Y   D   R   I   K   Q   D   L   M   F   R   I   V   K   K   K   L     611

3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT   3275
 612 K   D   P   E   F   V   I   R   K   Y   A   T   I   H   A   T   S   D   R   A     631

3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttatttttcattggaatttttaacaa   3343
 632 T   K   N   F   V   S   E   A   F   S   Y   F                                     643

3344 attctttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA   3405
 644               D   M   V   P   F   E   K   V   V   Q   L   L   S   M   K   T     659

3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT   3465
 660 S   D   T   L   F   V   D   F   V   D   Y   W   T   K   S   S   S   E   I   F     679

3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgttgaattgtaataaca   3532
 680 K   M   L   K   E   H   L   S   G   H   I   V   K                                 692

3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA     3593
 693               I   G   N   S   Q   Y   L   Q   K   V   G   I   P   Q   G   S     708

3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG   3653
 709 I   L   S   S   F   L   C   H   F   Y   M   E   D   L   I   D   E   Y   L   S     728

3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA   3713
 729 F   T   K   K   K   G   S   V   L   L   R   V   V   D   D   F   L   F   I   T     748

3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc   3777
 749 V   N   K   K   D   A   K   K   F   L   N   L   S   L   R   G                     764

3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA     3840
 765                       F   E   K   H   N   F   S   T   S   L   E   K   T   V     778

3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA   3900
 779 I   N   F   E   N   S   N   G   I   I   N   N   T   F   F   N   E   S   K   K     798
```

FIG. 15D

```
3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT  3960
 799 R   M   P   F   F   G   F   S   V   N   M   R   S   L   D   T   L   L   A   C    818

3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG  4020
 819 P   K   I   D   E   A   L   F   N   S   T   S   V   E   L   T   K   H   M   G    838

4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataatagctgacaaataatcag A TCG 4089
 839 K   S   F   F   Y   K   I   L   R                                           S    848

4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT  4149
 849 S   L   A   S   F   A   Q   V   F   I   D   I   T   H   N   S   K   F   N   S    868

4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA  4209
 869 C   C   N   I   Y   R   L   G   Y   S   M   C   M   R   A   Q   A   Y   L   K    888

4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
 889 R   M   K   D   I   F   I   P   Q   R   M   F   I   T   D                       903

4275 aaagtcattaattaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
 904                         L   L   N   V   I   G   R   K   I   W   K   K   L   A    917

4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
 918 E   I   L   G   Y   T   S   R   R   F   L   S   S   A   E   V   K   W           935

4402 ggtctcgagacttcagcaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA  4468
 936                                   L   F   C   L   G   M   R   D   G   L   K    946

4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT  4528
 947 P   S   F   K   Y   H   P   C   F   E   Q   L   I   Y   Q   F   Q   S   L   T    966

4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA AGA ATA  4588
 967 D   L   I   K   P   L   R   P   V   L   R   Q   V   L   F   L   H   R   R   I    986

4589 GCT GAT TAA tgtcattttcaatttattatatacatcctttattactggtgtcttaaacaatattattactaagtata 4665
 987 A   D   *                                                                       989
```

FIG. 15E

```
4666 gctgacccccaaagcaagcatactataggatttctagtaaagtaaaattaatctcgttattagttttgattgacttgtct 4745
4746 ttatccttatacttttaagaaagattgacagtggttgctgactactgcccacatgcccattaaacgggagtggttaaaca 4825
4826 ttaaaagtaatacatgaggctaatctcctttcatttagaataaggaaagtggttttctataatgaatgcccgcacta 4905
4906 atgcaaaaagacgaagattatcttctaaacaaggggattaagcatatccgaaggaaaagagagtaatatacccagtgtt 4985
4986 gttgaagaaagcaaggataatttggaacaagcttctgcagatgacaggctaaattttggtgaccgaattttggtaaaagc 5065
5066 cccaggttatccatggtggccggccttgctactgagacgaaaagaaactaaggatagtttgaatactaatagctcattta 5145
5146 atgtcttatataaggttttgtttttcctgacttcaattttgcatgggtgaaaagaaatagtgttaagccattattggat 5225
5226 tccgaaatagccaaatttcttggttcctcaaagcggaagtctaaagaacttattgaagcttatgaggcttcaaaaactcc 5305
5306 tcctgatttaaaggaggaatcttccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaatttttgc 5385
5386 aaaaaagaaaatatcattgggagacatctcttgatgaatcagatgcggagagtatctccagcggatccttgatgtcaata 5465
5466 acttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctctacgcagttaagtgaccaaaggtacc  5544
```

FIG. 15F

```
   1 gcagcgctgc gtcctgctgc gcacgtggga agccctggcc ccggccaccc ccgcgatgcc
  61 gcgcgctccc cgctgccgag ccgtgcgctc cctgctgcgc agccactacc gcgaggtgct
 121 gccgctggcc acgttcgtgc ggcgcctggg gccccagggc tggcggctgg tgcagcgcgg
 181 ggacccggcg gctttccgcg cgctggtggc ccagtgcctg gtgtgcgtgc cctgggacgc
 241 acggccgccc cccgccgccc cctccttccg ccaggtgtcc tgcctgaagg agctggtggc
 301 ccgagtgctg cagaggctgt gcgagcgcgg cgcgaagaac gtgctggcct tcggcttcgc
 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta
 421 cctgcccaac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgctgctgcg
 481 ccgcgtgggc gacgacgtgc tggttcacct gctggcacgc tgcgcgctct ttgtgctggt
 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgccgctg taccagctcg gcgctgccac
 601 tcaggcccgg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc
 661 ctggaaccat agcgtcaggg aggccggggt ccccctgggc ctgccagccc cgggtgcgag
 721 gaggcgcggg ggcagtgcca gccgaagtct gccgttgccc aagaggccca ggcgtggcgc
 781 tgcccctgag ccggagcgga cgcccgttgg gcaggggtcc tgggcccacc cgggcaggac
 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccg ccgaagaagc
 901 cacctctttg gagggtgcgc tctctggcac gcgccactcc cacccatccg tgggccgcca
 961 gcaccacgcg ggcccccat ccacatcgcg gccaccacgt ccctgggaca cgccttgtcc
1021 cccggtgtac gccgagacca agcacttcct ctactcctca ggcgacaagg agcagctgcg
1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga
1141 gaccatcttt ctgggttcca ggccctggat gccagggact ccccgcaggt tgccccgcct
1201 gccccagcgc tactggcaaa tgcggccccc gtttctggag ctgcttggga accacgcgca
1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgctg cgagctgcgg tcaccccagc
1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga
1381 cacagacccc cgtcgcctgg tgcagctgct ccgccagcac agcagcccct ggcaggtgta
1441 cggcttcgtg cgggcctgcc tgcgccggct ggtgccccca ggcctctggg gctccaggca
1501 caacgaacgc cgcttcctca ggaacaccaa gaagttcatc tccctgggga agcatgccaa
1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgcag
1621 gagcccaggg gttggctgtg ttccggccgc agagcaccgt ctgcgtgagg agatcctggc
1681 caagttcctg cactggctga tgagtgtgta cgtcgtcgag ctgctcaggt ctttctttta
1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag
1801 caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggagctgtc
1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg
1921 cttcatcccc aagcctgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc
1981 cagaacgttc cgcagagaaa agagggccga gcgtctcacc tcgagggtga aggcactgtt
2041 cagcgtgctc aactacgagc gggcgcggcg ccccggcctc ctgggcgcct ctgtgctggg
2101 cctggacgat atccacaggg cctggcgcac cttcgtgctg cgtgtgcggg cccaggaccc
2161 gccgcctgag ctgtactttg tcaaggtgga tgtgacgggc gcgtacgaca ccatccccca
2221 ggacaggctc acggaggtca tcgccagcat catcaaaccc cagaacacgt actgcgtgcg
2281 tcggtatgcc gtggtccaga aggccgccca tgggcacgtc cgcaaggcct tcaagagcca
2341 cgtctctacc ttgacagacc tccagccgta catgcgacag ttcgtggctc acctgcagga
2401 gaccagcccg ctgagggatg ccgtcgtcat cgagcagagc tcctccctga atgaggccag
2461 cagtggcctc ttcgacgtct tcctacgctt catgtgccac cacgccgtgc gcatcagggg
2521 caagtcctac gtccagtgcc aggggatccc gcagggctcc atcctctcca cgctgctctg
2581 cagcctgtgc tacggcgaca tggagaacaa gctgtttgcg gggattcggc gggacgggct
2641 gctcctgcgt ttggtggatg atttcttgtt ggtgacacct cacctcaccc acgcgaaaac
2701 cttcctcagg accctggtcc gaggtgtccc tgagtatggc tgcgtggtga acttgcggaa
2761 gacagtggtg aacttccctg tagaagacga ggccctgggt ggcacggctt ttgttcagat
2821 gccggcccac ggcctattcc cctggtgcgg cctgctgctg gatacccgga ccctggaggt
2881 gcagagcgac tactccagct atgcccggac ctccatcaga gccagtctca ccttcaaccg
2941 cggcttcaag gctgggagga acatgcgtcg caaactcttt ggggtcttgc ggctgaagtg
3001 tcacagcctg tttctggatt tgcaggtgaa cagcctccag acggtgtgca ccaacatcta
3061 caagatcctc ctgctgcagg cgtacaggtt tcacgcatgt gtgctgcagc tcccatttca
3121 tcagcaagtt tggaagaacc ccacattttt cctgcgcgtc atctctgaca cggcctccct
3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctgggggcca agggcgccgc
3241 cggccctctg ccctccgagg ccgtgcagtg gctgtgccac caagcattcc tgctcaagct
3301 gactcgacac cgtgtcacct acgtgccact cctggggtca ctcaggacag cccagacgca
3361 gctgagtcgg aagctcccgg ggacgacgct gactgccctg gaggccgcag ccaacccggc
3421 actgccctca gacttcaaga ccatcctgga ctgatggcca cccgcccaca gccaggccga
3481 gagcagacac cagcagccct gtcacgccgg gctctacgtc ccagggaggg aggggcggcc
3541 cacacccagg cccgcaccgc tgggagtctg aggcctgagt gagtgtttgg ccgaggcctg
3601 catgtccggc tgaaggctga gtgtccggct gaggcctgag cgagtgtcca gccaagggct
3661 gagtgtccag cacacctgcc gtcttcactt ccccacaggc tggcgctcgg ctccacccca
3721 gggccagctt ttcctcacca ggagcccggc ttccactccc cacataggaa tagtccatcc
3781 ccagattcgc cattgttcac ccctcgccct gccctccttt gccttccacc cccaccatcc
3841 aggtggagac cctgagaagg accctgggag ctctgggaat ttggagtgac caaaggtgtg
3901 ccctgtacac aggcgaggac cctgcacctg gatgggggtc cctgtgggtc aaattggggg
3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa
```

FIG. 16

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPQGWRLVQRGDP
AAFRALVAQCLVCVPWDARPPPAAPSFRQVSCLKELVARVLQRL
CERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALR
GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY
QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG
ARRRGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRG
PSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP
STSRPPRPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLLSSLRP
SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL
LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPQGSVAAPEE
EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE
RRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC
VPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTETTFQKNR
LFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPAL
LTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKA
LFSVLNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPP
ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ
KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPQGSI
LSTLLCSLCYGDMENKLFAGIRRDGLLLRLVDDFLLVTPHLTHA
KTFLRTLVRGVPEYGCVVNLRKTVVNFPVEDEALGGTAFVQMPA
HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
NMRRKLFGVLRLKCHSLFLDLQVNSLQTVCTNIYKILLLQAYRF
HACVLQLPFHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL
GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGSLRTAQ
TQLSRKLPGTTLTALEAAANPALPSDFKTILD

FIG. 17

GGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTT
TTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTG
GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT
GTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACT
CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGG
AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACT
GTTCAGCGTGCTCAACTACGAGCGGGCGCGGCGCCCCGGCCTCCTGGGCGCCTCTGTGCT
GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGA
CCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
CCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCCAGAACACGTACTGCGT
GCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAG
CCACGTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT
GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGACGGGC
TGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAA
CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACTTGCGGA
AGACAGTGGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA
TGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGG
TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC
GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT
GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT
ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTC
ATCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGCCTCCC
TCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGCCAAGGGCGCCG
CCGGCC7TCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGC
TGACTCGACACCGTGTCACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGC
AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGG
CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCG
AGAGCAGACACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGC
CCACACCCAGGCCTGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCT
GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC
TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
AGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC
CCCAGATTCGCCATTGTTCACCCCTCGCCCTGCCCTCCTTTGCCTTCCACCCCCACCATC
CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT
GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGG
GGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTG0AAAAAAAAAA
AAAAAAAAAAAAAAAA

FIG. 18

MetSerValTyrValValGluLeuLeuArgSerPhePhe
TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe
PheTyrArgLysSerValTrpSerLysLeuGlnSerIle
GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu
LeuSerGluAlaGluValArgGlnHisArgGluAlaArg
ProAlaLeuLeuThrSerArgLeuArgPheIleProLys
ProAspGlyLeuArgProIleValAsnMetAspTyrVal
ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu
ArgLeuThrSerArgValLysAlaLeuPheSerValLeu
AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla
SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg
ThrPheValLeuArgValArgAlaGlnAspProProPro
GluLeuTyrPheValLysValAspValThrGlyAlaTyr
AspThrIleProGlnAspArgLeuThrGluValIleAla
SerIleIleLysProGlnAsnThrTyrCysValArgArg
TyrAlaValValGlnLysAlaAlaHisGlyHisValArg
LysAlaPheLysSerHisValLeuArgProValProGly
AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln
ProValLeuArgArgHisGlyGluGlnAlaValCysGly
AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

FIG. 19

```
                                                                    1
                                                                   met
GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG

                                        10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

            20                                              30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

                                        40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

            50                                              60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

                                        70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

            80                                              90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

                                       100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

           110                                             120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC
```

FIG. 20A

```
                                        130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

            140                                         150
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

                                        160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

            170                                         180
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

                                        190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

            200                                         210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

                                        220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

            230                                         240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

                                        250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

            260                                         270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT


                                        280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

            290                                         300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

                                        310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

            320                                         330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC
```

FIG. 20B

```
                                        340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

            350                                         360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

                                    370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

            380                                         390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

                                    400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

            410                                         420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

                                    430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

            440                                         450
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

                                    460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

            470                                         480
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

                                    490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

            500                                         510
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

                                    520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

            530                                         540
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG
```

FIG. 20C

```
                                    550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

            560                                         570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

                                580
arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC
            590                                         600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

                                610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

            620                                         630
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

                                640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

            650                                         660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

                                670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

            680                                         690
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

                                700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

            710                                         720
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

                                730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

            740                                         750
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC
```

FIG. 20D

```
                                    760
his gly his val arg lys ala phe lys ser his val leu arg pro
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

            770                                         780
val pro gly asp pro ala gly leu his pro leu his ala ala leu
GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

                                790
gln pro val leu arg arg his gly glu gln ala val cys gly asp
CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

            800                             807
ser ala gly arg ala ala pro ala phe gly gly OP
TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT

GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA

GTATGGCTGCGTGGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTAGAAGACGAGGC

CCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCT

GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC

CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA

ACTCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTTTCTGGATTTGCAGGTGAACAG

CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA

CGCATGTGTGCTGCAGCTCCCATTTCATCAGCAAGTTTGGAAGAACCCCACATTTTTCCT

GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG

GATGTCGCTGGGGGCCAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCT

GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCT

GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC

TGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG

ATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTCACGCCGGGCT

CTACGTCCCAGGGAGGGAGGGGCGGCCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG

CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG

GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC

CACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC

CACTCCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCCTCGCCCTGCC

CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC

TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT

GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATG

AGTTTTTCAGTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 20E

```
   1 CCATGGGACCCACTGCAGGGGCAGCTGGGAGGCTGCAGGCTTCAGGTCCCAGTGGGGTTG
     GGTACCCTGGGTGACGTCCCCGTCGACCCTCCGACGTCCGAAGTCCAGGGTCACCCCAAC

  61 CCATCTGCCAGTAGAAACCTGATGTAGAATCAGGGCGCGAGTGTGGACACTGTCCTGAAT
     GGTAGACGGTCATCTTTGGACTACATCTTAGTCCCGCGCTCACACCTGTGACAGGACTTA

 121 CTCAATGTCTCAGTGTGTGCTGAAACATGTAGAAATTAAAGTCCATCCCTCCTACTCTAC
     GAGTTACAGAGTCACACACGACTTTGTACATCTTTAATTTCAGGTAGGGAGGATGAGATG

 181 TGGGATTGAGCCCCTTCCCTATCCCCCCCCAGGGGCAGAGGAGTTCCTCTCACTCCTGTG
     ACCCTAACTCGGGGAAGGGATAGGGGGGGGTCCCCGTCTCCTCAAGGAGAGTGAGGACAC

 241 GAGGAAGGAATGATACTTTGTTATTTTTCACTGCTGGTACTGAATCCACTGTTTCATTTG
     CTCCTTCCTTACTATGAAACAATAAAAAGTGACGACCATGACTTAGGTGACAAAGTAAAC

                          ****************************************
 301 TTGGTTTGTTTGTTTTGTTTTGAGAGGCGGTTTCACTCTTGTTGCTCAGGCTGGAGGGAG
     AACCAAACAAACAAAACAAAACTCTCCGCCAAAGTGAGAACAACGAGTCCGACCTCCCTC

     ************************************************************
 361 TGCAATGGCGCGATCTTGGCTTACTGCAGCCTCTGCCTCCCAGGTTCAAGTGATTCTCCT
     ACGTTACCGCGCTAGAACCGAATGACGTCGGAGACGGAGGGTCCAAGTTCACTAAGAGGA

                                                          alu
     ************************************************************
 421 GCTTCCGCCTCCCATTTGGCTGGGATTACAGGCACCCGCCACCATGCCCAGCTAATTTTT
     CGAAGGCGGAGGGTAAACCGACCCTAATGTCCGTGGGCGGTGGTACGGGTCGATTAAAAA

                                                             ==
     ************************************************************
 481 TGTATTTTTAGTAGAGACGGGGGTGGGGGTGGGGTTCACCATGTTGGCCAGGCTGGTCTC
     ACATAAAAATCATCTCTGCCCCCACCCCCACCCCAAGTGGTACAACCGGTCCGACCAGAG

          CAP
     =============>
     ************************************************************
 541 GAACTTCTGACCTCAGATGATCCACCTGCCTCTGCCTCCTAAAGTGCTGGGATTACAGGT
      CTTGAAGACTGGAGTCTACTAGGTGGACGGAGACGGAGGATTTCACGACCCTAATGTCCA

     **********************
 601 GTGAGCCACCATGCCCAGCTCAGAATTTACTCTGTTTAGAAACATCTGGGTCTGAGGTAG
     CACTCGGTGGTACGGGTCGAGTCTTAAATGAGACAAATCTTTGTAGACCCAGACTCCATC

                                          CCAAT
                                     ***************>
 661 GAAGCTCACCCCACTCAAGTGTTGTGGTGTTTTAAGCCAATGATAGAATTTTTTTATTGT
     CTTCGAGTGGGGTGAGTTCACAACACCACAAAATTCGGTTACTATCTTAAAAAAATAACA

 721 TGTTAGAACACTCTTGATGTTTTACACTGTGATGACTAAGACATCATCAGCTTTTCAAAG
     ACAATCTTGTGAGAACTACAAAATGTGACACTACTGATTCTGTAGTAGTCGAAAAGTTTC
```

FIG. 21A

```
                                        CAP
                                ***************>
781 ACACACTAACTGCACCCATAATACTGGGGTGTCTTCTGGGTATCAGCGATCTTCATTGAA
    TGTGTGATTGACGTGGGTATTATGACCCCACAGAAGACCCATAGTCGCTAGAAGTAACTT

                                                              CAP
                                                     ************
841 TGCCGGGAGGCGTTTCCTCGCCATGCACATGGTGTTAATTACTCCAGCATAATCTTCTGC
    ACGGCCCTCCGCAAAGGAGCGGTACGTGTACCACAATTAATGAGGTCGTATTAGAAGACG

    ***>
901 TTCCATTTCTTCTCTTCCCTCTTTTAAAATTGTGTTTTCTATGTTGGCTTCTCTGCAGAG
    AAGGTAAAGAAGAGAAGGGAGAAAATTTTAACACAAAAGATACAACCGAAGAGACGTCTC

                                CAP
                        ***************>
961 AACCAGTGTAAGCTACAACTTAACTTTTGTTGGAACAAATTTTCCAAACCGCCCCTTTGC
    TTGGTCACATTCGATGTTGAATTGAAAACAACCTTGTTTAAAAGGTTTGGCGGGGAAACG

1021 CCTAGTGGCAGAGACAATTCACAAACACAGCCCTTTAAAAAGGCTTAGGGATCACTAAGG
     GGATCACCGTCTCTGTTAAGTGTTTGTGTCGGGAAATTTTTCCGAATCCCTAGTGATTCC

1081 GGATTTCTAGAAGAGCGACCCGTAATCCTTAAGTATTTACAAGACGAGGCTAACCTCCAG
     CCTAAAGATCTTCTCGCTGGGCATTAGGAATTCATAAATGTTCTGCTCCGATTGGAGGTC

1141 CGAGCGTGACAGCCCAGGGAGGGTGCGAGGCCTGTTCAAATGCTAAGCTTCCATAAATAA
     GCTCGCACTGTCGGGTCCCTCCCACGCTCCGGACAAGTTTACGATTCGAAGGTATTTATT

1201 AGCAAATTTCCTCCGGCAGTTTCTGGAAAGTAGGAAAGGTTAACATTTAAGGTTGCGTTT
     TCGTTTAAAGGAGGCCGTCAAAGACCTTTCATCCTTTCCAATTGTAAATTCCAACGCAAA

1261 GTTAGCATTTCAGTGTTTGCCGACCTCAGCTAACAGCATCCCTGCAAGGCCTCGGGAGAC
     CAATCGTAAAGTCACAAACGGCTGGAGTCGATTGTCGTAGGGACGTTCCGGAGCCCTCTG

1321 CCAGAAGTTTCTCGCCCCTTAGATCCAAACTTGAGCAACCCGGAGTCTGGATTCCTGGGA
     GGTCTTCAAAGAGCGGGGAATCTAGGTTTGAACTCGTTGGGCCTCAGACCTAAGGACCCT

                        TopoII
                ******************>
1381 AGTCCTCAGCTGTCCTGCGGTTGTGCCGGGGCCCCAGGTCTGGAGGGGACCAGTGGCCGT
     TCAGGAGTCGACAGGACGCCAACACGGCCCCGGGGTCCAGACCTCCCCTGGTCACCGGCA

1441 GTGGCTTCTACTGCTGGGCTGGAAGTCGGGCCTCCTAGCTCTGCAGTCCGAGGCTTGGAG
     CACCGAAGATGACGACCCGACCTTCAGCCCGGAGGATCGAGACGTCAGGCTCCGAACCTC

1501 CCAGGTGCCTGGACCCCGAGGCTGCCCTCCACCCTGTGCGGGCGGGATGTGACCAGATGT
     GGTCCACGGACCTGGGGCTCCGACGGGAGGTGGGACACGCCCGCCCTACACTGGTCTACA

1561 TGGCCTCATCTGCCAGACAGAGTGCCGGGGCCCAGGGTCAAGGCCGTTGTGGCTGGTGTG
     ACCGGAGTAGACGGTCTGTCTCACGGCCCCGGGTCCCAGTTCCGGCAACACCGACCACAC

1621 AGGCGCCCGGTGCGCGGCCAGCAGGAGCGCCTGGCTCCATTTCCCACCCTTTCTCGACGG
     TCCGCGGGCCACGCGCCGGTCGTCCTCGCGGACCGAGGTAAAGGGTGGGAAAGAGCTGCC
```

FIG. 21B

```
1681 GACCGCCCCGGTGGGTGATTAACAGATATTGGGGTGGTTTGCTCATGGTGGGGACCCCTT
     CTGGCGGGGCCACCCACTAATTGTCTATAACCCCACCAAACGAGTACCACCCCTGGGGAA

1741 CGCCGCCTGAGAACCTGCAAAGAGAAATGACGGGCCTGTGTCAAGGAGCCCAAGTCGCGG
     GCGGCGGACTCTTGGACGTTTCTCTTTACTGCCCGGACACAGTTCCTCGGGTTCAGCGCC

1801 GGAAGTGTTGCAGGGAGGCACTCCGGGAGGTCCCGCGTGCCCGTCCAGGGAGCAATGCGT
     CCTTCACAACGTCCCTCCGTGAGGCCCTCCAGGGCGCACGGGCAGGTCCCTCGTTACGCA

1861 CCTCGGGTTCGTCCCCAGCCGCGTCTACGCGCCTCCGTCCTCCCCTTCACGTCCGGCATT
     GGAGCCCAAGCAGGGGTCGGCGCAGATGCGCGGAGGCAGGAGGGGAAGTGCAGGCCGTAA

1921 CGTGGTGCCCGGAGCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGA
     GCACCACGGGCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTCGGGACCCAGAGGCCT

1981 TCAGGCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCCAGGGCCTCCACATCATGGC
     AGTCCGGTCGCCGGTTTCCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCG

2041 CCCTCCCTCGGGTTACCCCACAGCCTAGGCCGATTCGACCTCTCTCCGCTGGGGCCCTCG
     GGGAGGGAGCCCAATGGGGTGTCGGATCCGGCTAAGCTGGAGAGAGGCGACCCCGGGAGC

                                                Sp1
                                          ********
2101 CTGGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGCGGGCGGGGAAGCGCGGCCCAGAC
     GACCGCAGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCCTTCGCGCCGGGTCTG

2161 CCCCGGGTCCGCCCGGAGCAGCTGCGCTGTCGGGGCCAGGCCGGGCTCCCAGTGGATTCG
     GGGGCCCAGGCGGGCCTCGTCGACGCGACAGCCCCGGTCCGGCCCGAGGGTCACCTAAGC

2221 CGGGCAACAGACGCCCAGGACCGCGCTTCCCACGTGGCGGAGGGACTGGGGACCCGGGCA
     GCCCGTTGTCTGCGGGTCCTGGCGCGAAGGGTGCACCGCCTCCCTGACCCCTGGGCCCGT

                                                         Sp1
                                                      ========
                                                  E2F
                                             ********
2281 CCGGTCCTGCCCCTTCACCTTCCAGCTCCGCCTCGTCCGCGCGGAACCCCGCCCCGTCCC
     GGCCAGGACGGGGAAGTGGAAGGTCGAGGCGGAGCAGGCGCGCCTTGGGGCGGGGCAGGG

2341 GAACCCTTCCCGGGTCCCCGGCCCAGCCCCTTCCGGGCCATCCCAGCCCGTCCCGTTCCT
     CTTGGGAAGGGCCCAGGGGCCGGGTCGGGGAAGGCCCGGTAGGGTCGGGCAGGGCAAGGA

                 Sp1
              ========
        E2F                         NFkB
     *********                     *********  *********************
2401 TTTCCGCGGCCCCGCCCTCTCCTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCCTGCTGC
     AAAGGCGCCGGGGCGGGAGAGGAGCGCCGCGCTCAAAGTCCGTCGCGACGCAGGACGACG

       hTRT5'
     ******************************>
2461 GCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCGATGCCGCGCGCTCCCCGCTGCCGAG
     CGTGCACCCTTCGGGACCGGGGCCGGTGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTC

2521 CCGTGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTCGTGC
     GGCACGCGAGGGACGACGCGTCGGTGATGGCGCTCCACGACGGCGACCGGTGCAAGCACG
```

FIG. 21C

```
                                                                  E2F
                                                                *******
2581 GGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCCGCG
     CCGCGGACCCCGGGGTCCCGACCGCCGACCACGTCGCGCCCCTGGGCCGCCGAAAGGCGC

     *
2641 CGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCCGCCGCCC
     GCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCGTGCCGGCGGGGGGCGGCGGG

                         NFkB
                       ==========

                     **********************************************
2701 CCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCG
     GGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGCCGACCCCAACTCCCGCCGGC

                                                    Topo_II_cleavag
                                                     ::::::::::::::
                                                    NFkB
                                                  ++++++++++
                                                    NFkB
                                                  ===========
   Intron1
   ***********************************************************>
2761 GGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGT
     CCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAGTCCCGCGAAGGGGGCGTCCA

     e_site
     ::::
2821 GTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGCGAA
     CAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTT

2881 GAACGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCGCGGGGGCCCCCCCCGAGGC
     CTTGCACGACCGGAAGCCGAAGCGCGACGACCTGCCCCGGGCGCCCCCGGGGGGGCTCCG

2941 CTTCACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGAG
     GAAGTGGTGGTCGCACGCGTCGATGGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTC

3001 CGGGGCGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGC
     GCCCCGCACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCG

3061 ACGCTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCC
     TGCGACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGG

3121 GCTGTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCG
     CGACATGGTCGAGCCGCGACGGTGAGTCCGGGCCGGGGGCGGTGTGCGATCACCTGGGGC

3181 AAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCT
     TTCCGCAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGA

3241 GGGCCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTT
     CCCGGACGGTCGGGGCCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAA

3301 GCCCAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGG
     CGGGTTCTCCGGGTCCGCACCGCGACGGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCC
```

FIG. 21D

```
3361 GTCCTGGGCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGGTGTC
     CAGGACCCGGGTGGGCCCGTCCTGCGCACCTGGCTCACTGGCACCAAAGACACACCACAG

3421 ACCTGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCA
     TGGACGGTCTGGGCGGCTTCTTCGGTGGAGAAACCTCCCACGCGAGAGACCGTGCGCGGT

3481 CTCCCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCCATCCACATCGCGGCCACC
     GAGGGTGGGTAGGCACCCGGCGGTCGTGGTGCGCCCGGGGGGTAGGTGTAGCGCCGGTGG

3541 ACGTCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTC
     TGCAGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAG

3601 CTCAGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCT
     GAGTCCGCTGTTCCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTCGGA

3661 GACTGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCCTGGATGCCAGG
     CTGACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCC

3721 GACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCCTGTTTCT
     CTGAGGGGCGTCCAACGGGGCGGACGGGGTCGCGATGACCGTTTACGCCGGGGACAAAGA

3781 GGAGCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAAGACGCACTGCCC
     CCTCGACGAACCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGG

3841 GCTGCGAGCTGCGGTCACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTC
     CGACGCTCGACGCCAGTGGGGTCGTCGGCCACAGACACGGGCCCTCTTCGGGGTCCCGAG

3901 TGTGGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCA
     ACACCGCCGGGGGCTCCTCCTCCTGTGTCTGGGGGCAGCGGACCACGTCGACGAGGCGGT

3961 GCACAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCC
     CGTGTCGTCGGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGG

4021 CCCAGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTT
     GGGTCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTGGTTCTTCAA

4081 CATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGT
     GTAGAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCA

                                        ****************************
4141 GCGGGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCC
     CGCCCTGACGCGAACCGACGCGTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGG

                                                      Intron2
     ************************************************************
4201 AGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCCT
     TCCGGGGTCTCGACTTACGTCATCCCCGAGTCTTTTCCCCCGTCCGTCTCGGGACCAGGA

     ************************************************************
4261 CCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACG
     GGACAGAGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGC

     *****>
4321 GTGATCGAGGTCGAC
     CACTAGCTCCAGCTG
```

FIG. 21E

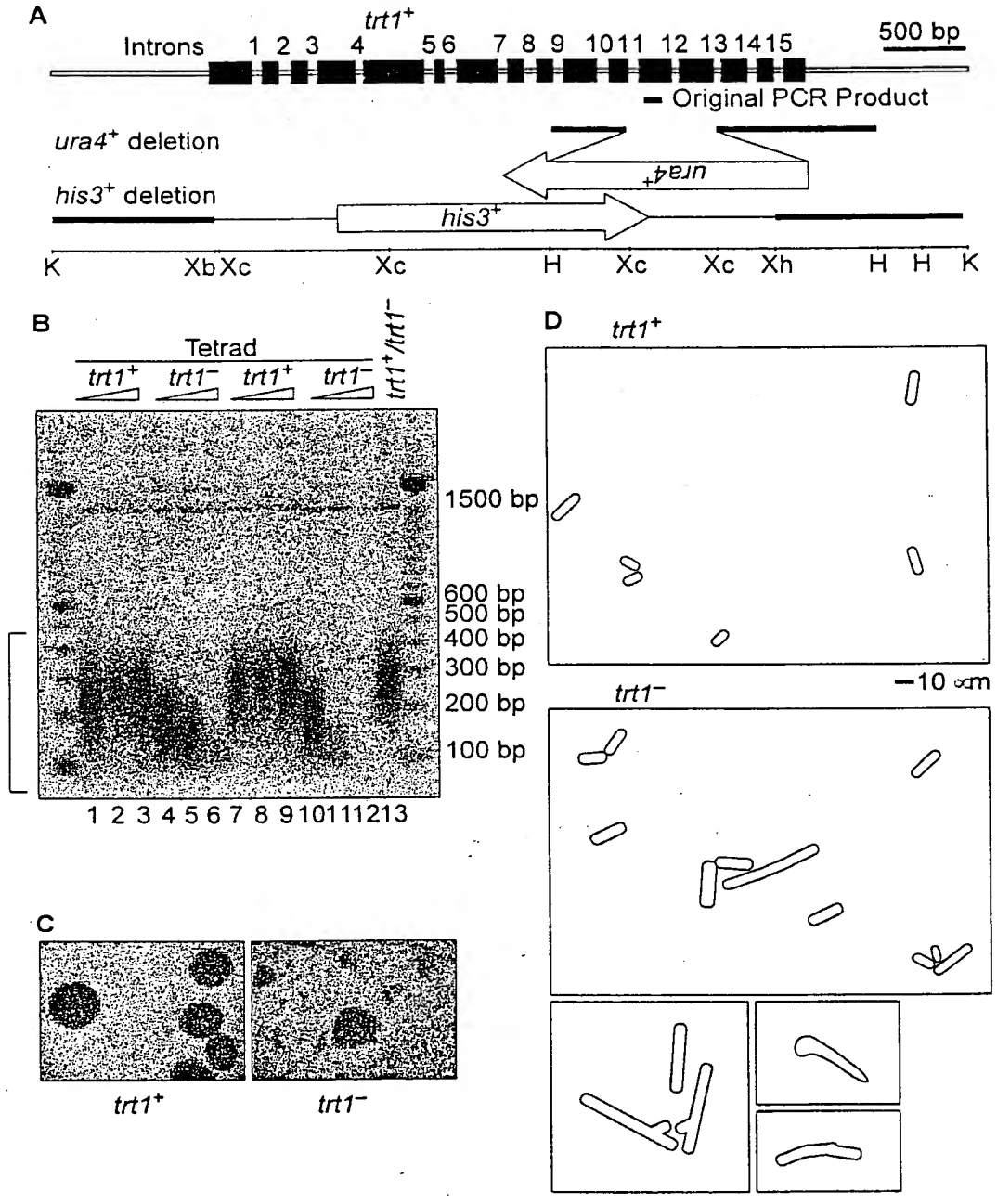


FIG. 22

gccaagttcctgcactggctgatgagtgtgtacgtcgtcgagctgctcaggtctttcttt
tatgtcacggagaccacgtttcaaaagaacaggctctttttctaccggaagagtgtctgg
agcaagttgcaaagcattggaatcagacagcacttgaagagggtgcagctgcgggacgtg
tcggaagcagaggtcaggcagcatcgggaagccaggcccgccctgctgacgtccagactc
cgcttcatccccaagcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
gccagaacgttccgcagagaaaagagggccgagcgtctcacctcgagggtgaaggcactg
ttcagcgtgctcaactacgagcgggcgcg

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
ACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC
AGGGGCAAGTC

FIG. 24

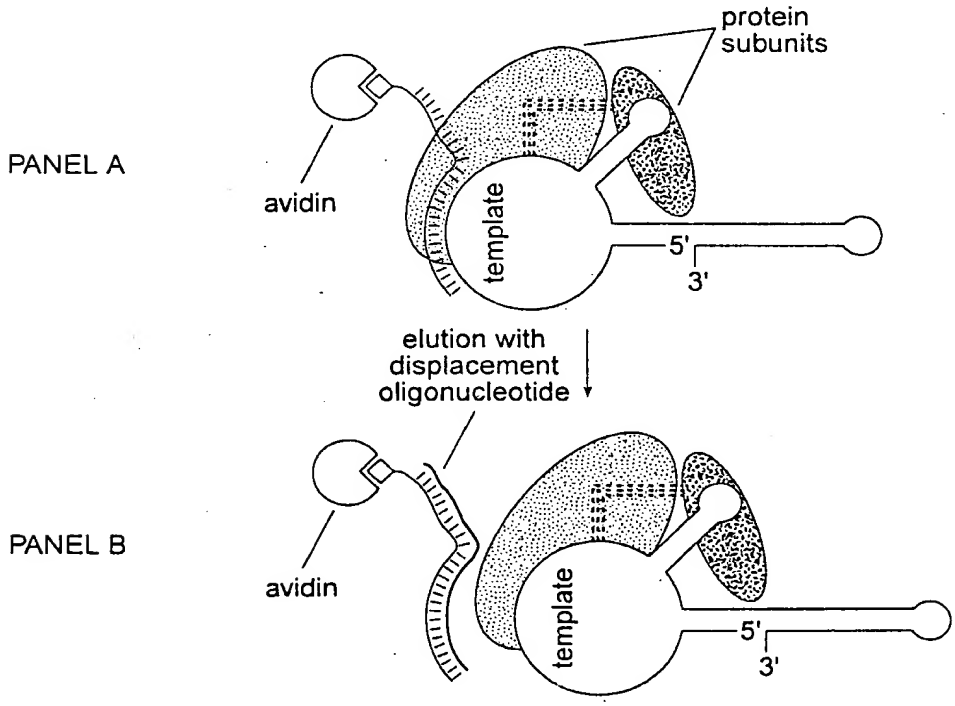


FIG. 26

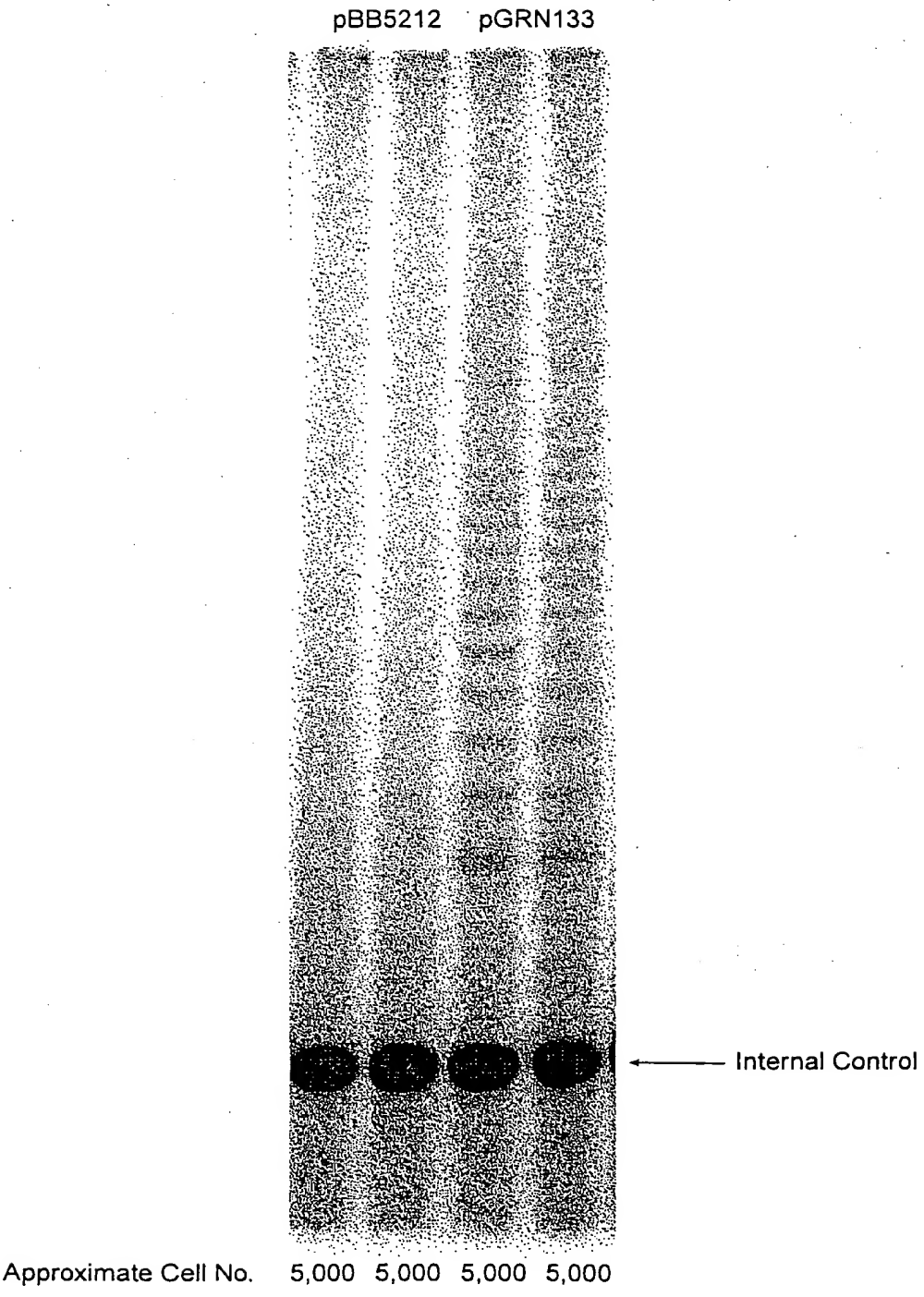


FIG. 25

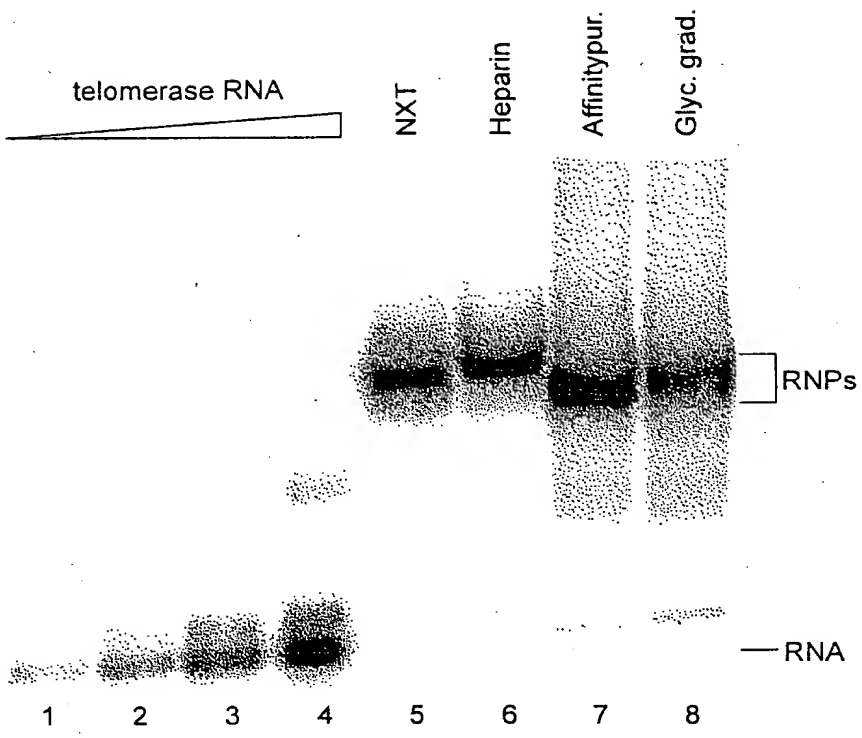


FIG. 27

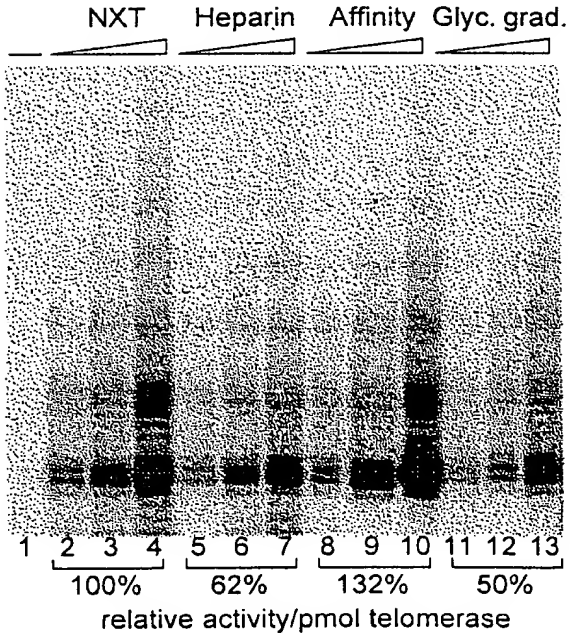


FIG. 28

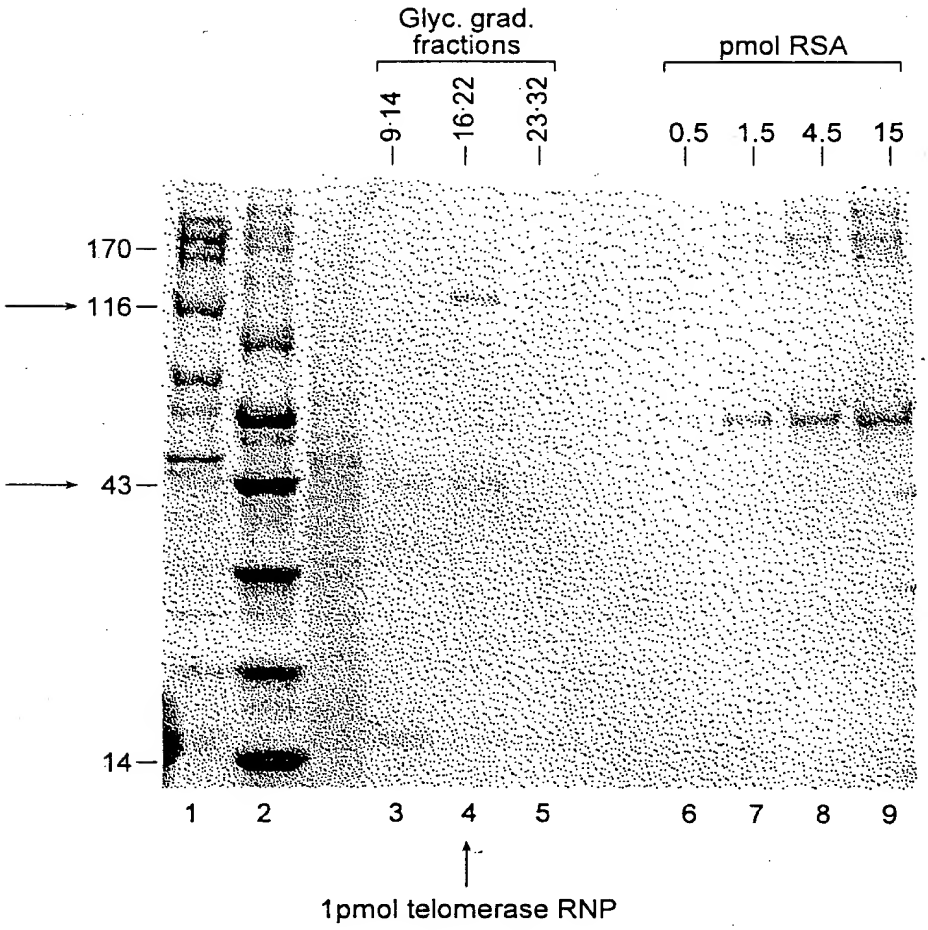


FIG. 29

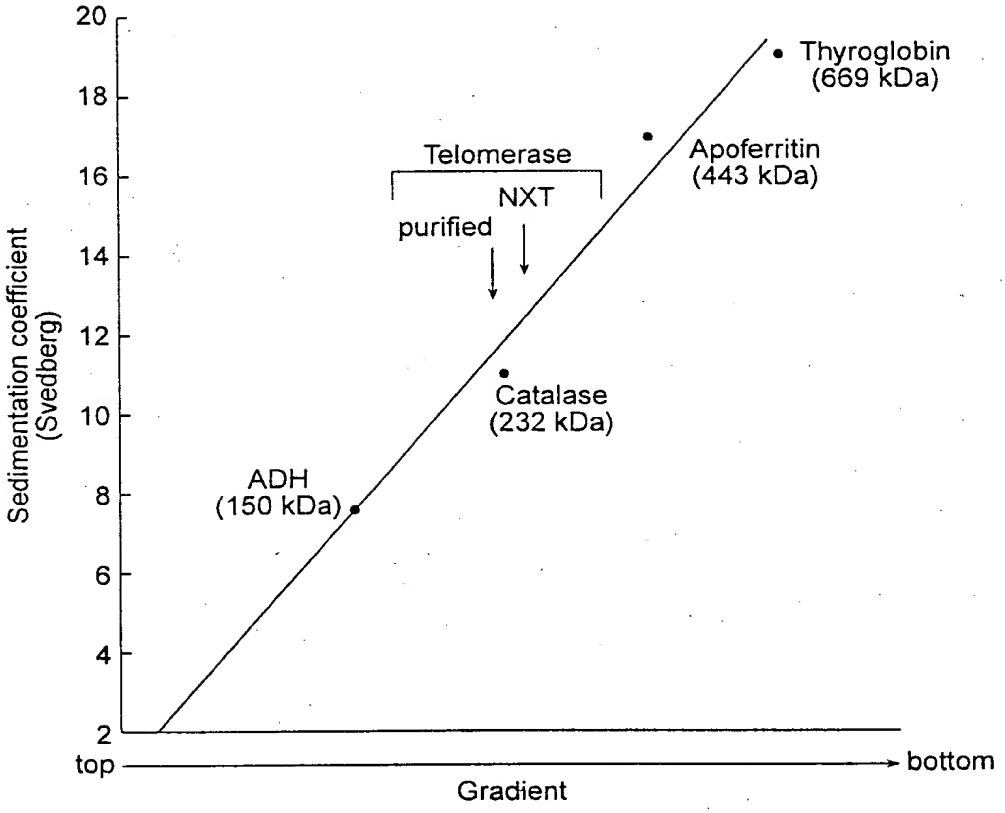


FIG. 30

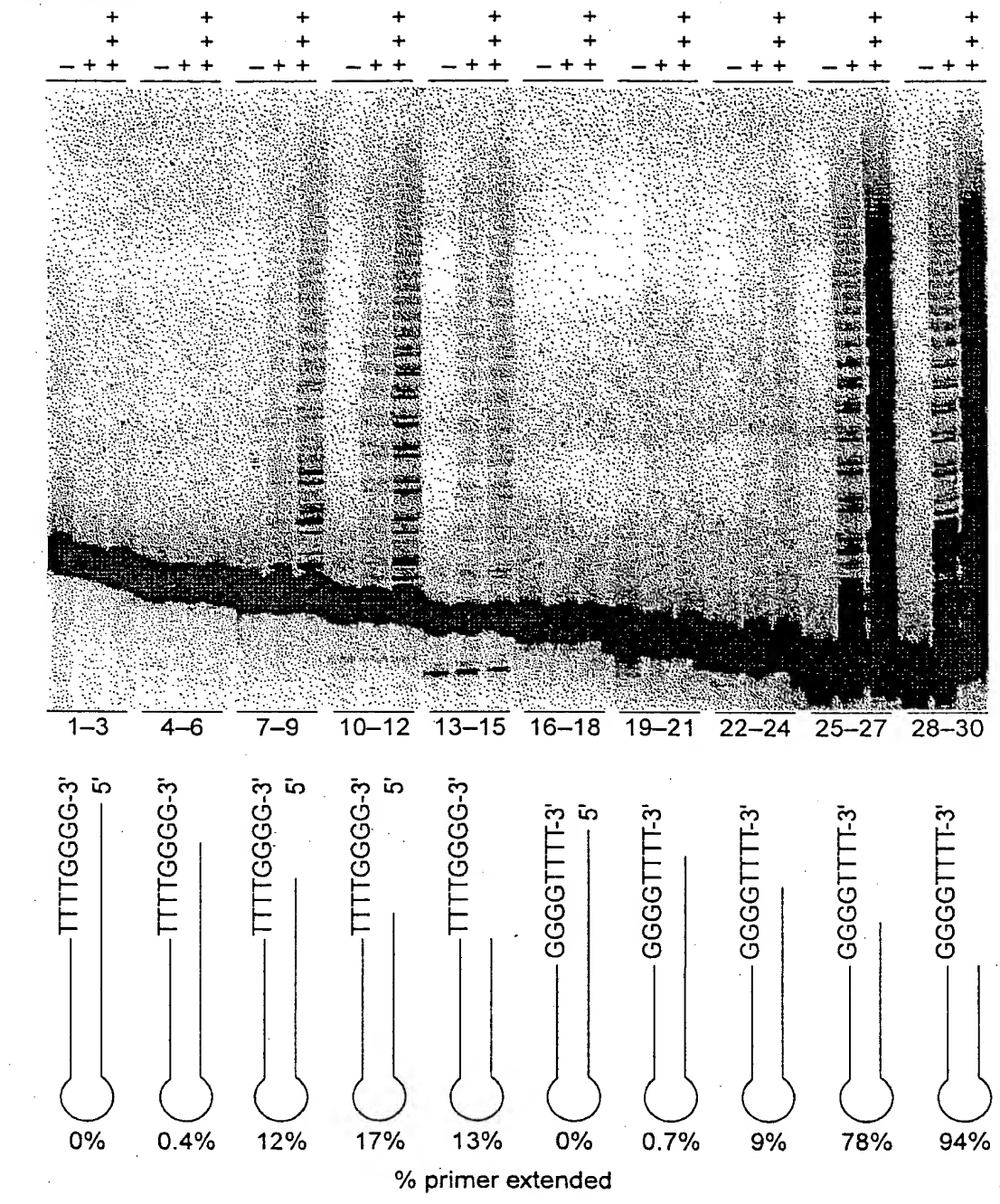

Telomerase:
1–3
4–6
7–9
10–12
13–15
16–18
19–21
22–24
25–27
28–30
TTTTGGGG-3'
5'
GGGGTTTT-3'
0%
0.4%
12%
17%
13%
0%
0.7%
9%
78%
94%
% primer extended

FIG. 31

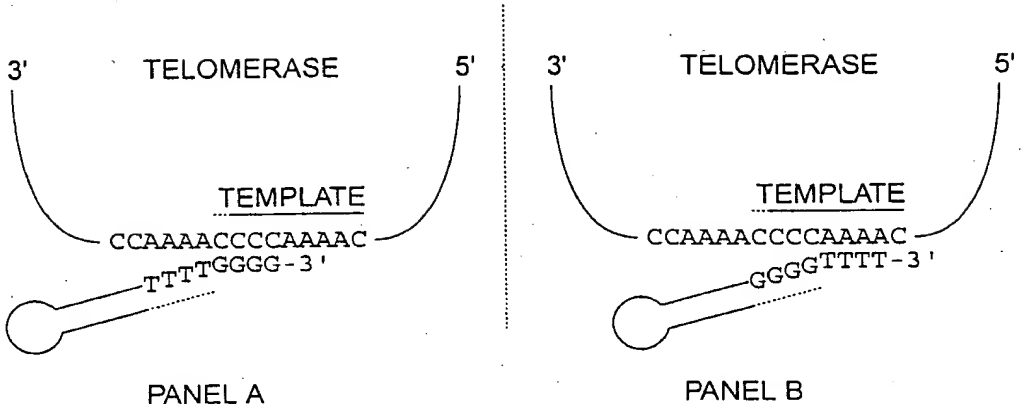


FIG. 32

```
   1  CCCCAAAACC CCAAAACCCC AAAACCCCTA TAAAAAAAGA AAAAATTGAG
  51  GTAGTTTAGA AATAAAATAT TATTCCCGCA CAAATGGAGA TGGATATTGA
 101  TTTGGATGAT ATAGAAAATT TACTTCCTAA TACATTCAAC AAGTATAGCA
 151  GCTCTTGTAG TGACAAGAAA GGATGCAAAA CATTGAAATC TGGCTCGAAA
 201  TCGCCTTCAT TGACTATTCC AAAGTTGCAA AAACAATTAG AGTTCTACTT
 251  CTCGGATGCA AATCTTTATA ACGATTCTTT CTTGAGAAAA TTAGTTTTAA
 301  AAAGCGGAGA GCAAAGAGTA GAAATTGAAA CATTACTAAT GTTTAAATAA
 351  AATCAGGTAA TGAGGATTAT TCTATTTTTT AGATCACTTC TTAAGGAGCA
 401  TTATGGAGAA AATTACTTAA TACTAAAAGG TAAACAGTTT GGATTATTTC
 451  CCTAGCCAAC AATGATGAGT ATATTAAATT CATATGAGAA TGAGTCAAAG
 501  GATCTCGATA CATCAGACTT ACCAAAGACA AACTCGCTAT AAAACGCAAG
 551  AAAAAGTTTG ATAATCGAAC AGCAGAAGAA CTTATTGCAT TTACTATTCG
 601  TATGGGTTTT ATTACAATTG TTTTAGGTAT CGACGGTGAA CTCCCGAGTC
 651  TTGAGACAAT TGAAAAAGCT GTTTACAACT GAAGGAATCG CAGTTCTGAA
 701  AGTTCTGATG TGTATGCCAT TATTTTGTGA ATTAATCTCA AATATCTTAT
 751  CTCAATTTAA TGGATAGCTA TAGAAACAAA CCAAATAAAC CATGCAAGTT
 801  TAATGGAATA TACGTTAAAT CCTTTGGGAC AAATGCACAC TGAATTTATA
 851  TTGGATTCTT AAAGCATAGA TACACAGAAT GCTTTAGAGA CTGATTTAGC
 901  TTACAACAGA TTACCTGTTT TGATTACTCT TGCTCATCTC TTATATCTTT
 951  AAAAGAAGCA GGCGAAATGA AAAGAAGACT AAAGAAAGAG ATTTCAAAAT
1001  TTGTTGATTC TTCTGTAACC GGAATTAACA ACAAGAATAT TAGCAACGAA
1051  AAAGAAGAAG AGCTATCACA ATCCTGATTC TTAAAGATTT CAAAAATTCC
1101  AGGTAAGAGA GATACATTCA TTAAAATTCA TATATTATAG TTTTTCATTT
1151  CACAGCTGTT ATTTTCTTTT ATCTTAACAA TATTTTTTGA TTAGCTGGAA
1201  GTAAAAAGTA TCAAATAAGA GAAGCGCTAG ACTGAGGTAA CTTAGCTTAT
1251  TCACATTCAT AGATCGACCT TCATATATCC AATACGATGA TAAGGAAACA
1301  GCAGTCATCC GTTTTAAAAA TAGTGCTATG AGGACTAAAT TTTTAGAGTC
1351  AAGAAATGGA GCCGAAATCT TAATCAAAAA GAATTGCGTC GATATTGCAA
1401  AAGAATCGAA CTCTAAATCT TTCGTTAATA AGTATTACCA ATCTTGATTG
1451  ATTGAAGAGA TTGACGAGGC AACTGCACAG AAGATCATTA AAGAAATAAA
1501  GTAACTTTTA TTAATTAGAG AATAAACTAA ATTACTAATA TAGAGATCAG
1551  CGATCTTCAA TTGACGAAAT AAAAGCTGAA CTAAAGTTAG ACAATAAAAA
1601  ATACAAACCT TGGTCAAAAT ATTGAGGAAG GAAAAGAAGA CCAGTTAGCA
1651  AAAGAAAAAA TAAGGCAATA AATAAAATGA GTACAGAAGT GAAGAAATAA
1701  AAGATTTATT TTTTTCAATA ATTTATTGAA AAGAGGGGTT TTGGGGTTTT
1751  GGGGTTTTGG GG
```

FIG. 34

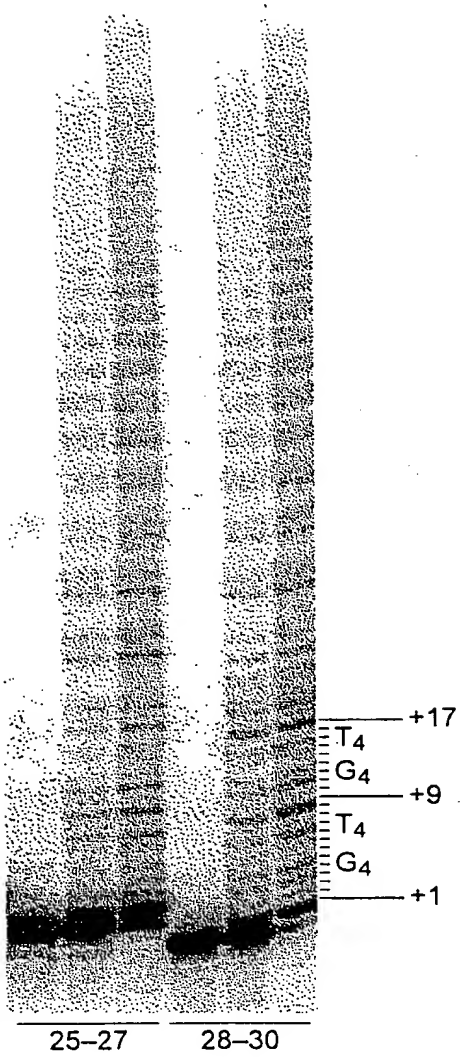


FIG. 33

```
    CCCCAAAACCCCAAAACCCCAAAACCCCTATAAAAAAAGAAAAAATTGAGGTAGTTTAGA
  1 ---------+---------+---------+---------+---------+---------+ 60
    GGGGTTTTGGGGTTTTGGGGTTTTGGGGATATTTTTTTCTTTTTTAACTCCATCAAATCT

a   P  Q  N  P  K  T  P  K  P  L  *  K  K  K  K  L  R  *  F  R  -
b    P  K  T  P  K  P  Q  N  P  Y  K  K  R  K  N  C  G  S  L  E  -
c     P  K  P  Q  N  P  K  T  P  I  K  K  E  K  I  E  V  V  *  K -

    AATAAAATATTATTCCCGCACAAATGGAGATGGATATTGATTTGGATGATATAGAAAATT
 61 ---------+---------+---------+---------+---------+---------+ 120
    TTATTTTATAATAAGGGCGTGTTTACCTCTACCTATAACTAAACCTACTATATCTTTTAA

a   N  K  I  L  F  P  H  K  W  R  W  I  L  I  W  M  I  *  K  I  -
b    I  K  Y  Y  S  R  T  N  G  D  G  Y  C  F  G  C  Y  R  K  F  -
c     *  N  I  I  P  A  Q  M  E  M  D  I  D  L  D  D  I  E  N  L -

    TACTTCCTAATACATTCAACAAGTATAGCAGCTCTTGTAGTGACAAGAAAGGATGCAAAA
121 ---------+---------+---------+---------+---------+---------+ 180
    ATGAAGGATTATGTAAGTTGTTCATATCGTCGAGAACATCACTGTTCTTTCCTACGTTTT

a   Y  F  L  I  H  S  T  S  I  A  A  L  V  V  T  R  K  D  A  K  -
b    T  S  *  Y  I  Q  Q  V  *  Q  L  L  *  C  Q  E  R  M  Q  N  -
c     L  P  N  T  F  N  K  Y  S  S  S  C  S  D  K  K  G  C  K  T -

    CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCCAAAGTTGCAAAAACAATTAG
181 ---------+---------+---------+---------+---------+---------+ 240
    GTAACTTTAGACCGAGCTTTAGCGGAAGTAACTGATAAGGTTTCAACGTTTTTGTTAATC

a   H  C  N  L  A  R  N  R  L  H  C  L  F  Q  S  C  K  N  N  *  -
b    I  E  I  W  L  E  I  A  F  I  D  Y  S  K  V  A  K  T  I  R  -
c     L  K  S  G  S  K  S  P  S  L  T  I  P  K  L  Q  K  Q  L  E -

    AGTTCTACTTCTCGGATGCAAATCTTTATAACGATTCTTTCTTGAGAAAATTAGTTTTAA
241 ---------+---------+---------+---------+---------+---------+ 300
    TCAAGATGAAGAGCCTACGTTTAGAAATATTGCTAAGAAAGAACTCTTTTAATCAAAATT

a   S  S  T  S  R  M  Q  I  F  I  T  I  L  S  C  E  N  *  F  *  -
b    V  L  L  L  G  C  K  S  L  *  R  F  F  L  E  K  I  S  F  K  -
c     F  Y  F  S  D  A  N  L  Y  N  D  S  F  L  R  K  L  V  L  K -

    AAAGCGGAGAGCAAAGAGTAGAAATTGAAACATTACTAATGTTTAAATAAAATCAGGTAA
301 ---------+---------+---------+---------+---------+---------+ 360
    TTTCGCCTCTCGTTTCTCATCTTTAACTTTGTAATGATTACAAATTTATTTTAGTCCATT

a   K  A  E  S  K  E  *  K  L  K  H  Y  *  C  L  N  K  I  R  *  -
b    K  R  R  A  K  S  R  N  C  N  I  T  N  V  *  I  K  S  G  N  -
c     S  G  E  Q  R  V  E  I  E  T  L  L  M  F  K  *  N  Q  V  M -

    TGAGGATTATTCTATTTTTTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA
361 ---------+---------+---------+---------+---------+---------+ 420
    ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCTCGTAATACCTCTTTTAATGAATT

a   C  G  L  F  Y  F  L  D  H  F  L  R  S  I  M  E  K  I  T  *  -
b    E  D  Y  S  I  F  *  I  T  S  *  G  A  L  W  R  K  L  L  N  -
c     R  I  I  L  F  F  R  S  L  L  K  E  H  Y  G  E  N  Y  L  I -
```

FIG. 35A

```
    TACTAAAAGGTAAACAGTTTGGATTATTTCCCTAGCCAACAATGATGAGTATATTAAATT
421 ---------+---------+---------+---------+---------+---------+ 480
    ATGATTTTCCATTTGTCAAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTTAA

a   Y  *  K  V  N  S  L  D  Y  F  P  S  Q  Q  C  C  V  Y  *  I  -
b    T  K  R  *  T  V  W  I  I  S  L  A  N  N  D  E  Y  I  K  F -
c     L  K  G  K  Q  F  G  L  F  P  *  P  T  M  M  S  I  L  N  S -

    CATATGAGAATGAGTCAAAGGATCTCGATACATCAGACTTACCAAAGACAAACTCGCTAT
481 ---------+---------+---------+---------+---------+---------+ 540
    GTATACTCTTACTCAGTTTCCTAGAGCTATGTAGTCTGAATGGTTTCTGTTTGAGCGATA

a   H  M  R  M  S  Q  R  I  S  I  H  Q  T  Y  Q  R  Q  T  R  Y  -
b    I  C  E  C  V  K  G  S  R  Y  I  R  L  T  K  D  K  L  A  I -
c     Y  E  N  E  S  K  D  L  D  T  S  D  L  P  K  T  N  S  L  * -

    AAAACGCAAGAAAAAGTTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTCG
541 ---------+---------+---------+---------+---------+---------+ 600
    TTTTGCGTTCTTTTTCAAACTATTAGCTTGTCGTCTTCTTGAATAACGTAAATGATAAGC

a   K  T  Q  E  K  V  C  *  S  N  S  R  R  T  Y  C  I  Y  Y  S  -
b    K  R  K  K  K  F  D  N  R  T  A  E  E  L  I  A  F  T  I  R -
c     N  A  R  K  S  L  I  I  E  Q  Q  K  N  L  L  H  L  L  F  V -

    TATGGGTTTTATTACAATTGTTTTAGGTATCGACGGTGAACTCCCGAGTCTTGAGACAAT
601 ---------+---------+---------+---------+---------+---------+ 660
    ATACCCAAAATAATGTTAACAAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA

a   Y  G  F  Y  Y  N  C  F  R  Y  R  R  C  T  P  E  S  C  D  N  -
b    M  G  F  I  T  I  V  L  G  I  D  G  E  L  P  S  L  E  T  I -
c     W  V  L  L  Q  L  F  *  V  S  T  V  N  S  R  V  L  R  Q  L -

    TGAAAAAGCTGTTTACAACTGAAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT
661 ---------+---------+---------+---------+---------+---------+ 720
    ACTTTTTCGACAAATGTTGACTTCCTTAGCGTCAAGACTTTCAAGACTACACATACGGTA

a   C  K  S  C  L  Q  L  K  E  S  Q  F  C  K  F  C  C  V  C  H  -
b    E  K  A  V  Y  N  C  R  N  R  S  S  E  S  S  D  V  Y  A  I -
c     K  K  L  F  T  T  E  G  I  A  V  L  K  V  L  M  C  M  P  L -

    TATTTTGTGAATTAATCTCAAATATCTTATCTCAATTTAATGGATAGCTATAGAAACAAA
721 ---------+---------+---------+---------+---------+---------+ 780
    ATAAAACACTTAATTAGAGTTTATAGAATAGAGTTAAATTACCTATCGATATCTTTGTTT

a   Y  F  V  N  *  S  Q  I  S  Y  L  N  L  M  D  S  Y  R  N  K  -
b    I  L  C  I  N  L  K  Y  L  I  S  I  *  W  I  A  I  E  T  N -
c     F  C  E  L  I  S  N  I  L  S  Q  F  N  G  *  L  *  K  Q  T -

    CCAAATAAACCATGCAAGTTTAATGGAATATACGTTAAATCCTTTGGGACAAATGCACAC
781 ---------+---------+---------+---------+---------+---------+ 840
    GGTTTATTTGGTACGTTCAAATTACCTTATATGCAATTTAGGAAACCCTGTTTACGTGTG

a   P  N  K  P  C  K  F  N  G  I  Y  V  K  S  F  G  T  N  A  H  -
b    Q  I  N  H  A  S  L  M  E  Y  T  L  N  P  L  G  Q  M  H  T -
c     K  *  T  M  Q  V  *  W  N  I  R  *  I  L  W  D  K  C  T  L -

    TGAATTTATATTGGATTCTTAAAGCATAGATACACAGAATGCTTTAGAGACTGATTTAGC
841 ---------+---------+---------+---------+---------+---------+ 900
    ACTTAAATATAACCTAAGAATTTCGTATCTATGTGTCTTACGAAATCTCTGACTAAATCG

a   C  I  Y  I  G  F  L  K  H  R  Y  T  E  C  F  R  D  C  F  S  -
b    E  F  I  L  D  S  *  S  I  D  T  Q  N  A  L  E  T  D  L  A -
c     N  L  Y  W  I  L  K  A  *  I  H  R  M  L  *  R  L  I  *  L -
```

FIG. 35B

```
     TTACAACAGATTACCTGTTTTGATTACTCTTGCTCATCTCTTATATCTTTAAAAGAAGCA
 901 ---------+---------+---------+---------+---------+---------+ 960
     AATGTTGTCTAATGGACAAAACTAATGAGAACGAGTAGAGAATATAGAAATTTTCTTCGT

a    L  Q  Q  I  T  C  F  D  Y  S  C  S  S  L  I  S  L  K  E  A   -
b     Y  N  R  L  P  V  L  I  T  L  A  H  L  L  Y  L  *  K  K  Q  -
c      T  T  D  Y  L  F  C  L  L  L  L  I  S  Y  I  F  K  R  S  R -

     GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTTCAAAATTTGTTGATTCTTCTGTAACC
 961 ---------+---------+---------+---------+---------+---------+ 1020
     CCGCTTTACTTTTCTTCTGATTTCTTTCTCTAAAGTTTTAAACAACTAAGAAGACATTGG

a    G  E  M  K  R  R  L  K  K  E  I  S  K  F  V  D  S  S  V  T   -
b     A  K  C  K  E  D  *  R  K  R  F  Q  N  L  L  I  L  L  *  P  -
c      R  N  E  K  K  T  K  E  R  D  F  K  I  C  C  F  F  C  N  R -

     GGAATTAACAACAAGAATATTAGCAACGAAAAAGAAGAAGAGCTATCACAATCCTGATTC
1021 ---------+---------+---------+---------+---------+---------+ 1080
     CCTTAATTGTTGTTCTTATAATCGTTGCTTTTTCTTCTTCTCGATAGTGTTAGGACTAAG

a    G  I  N  N  K  N  I  S  N  E  K  E  E  E  L  S  Q  S  C  F   -
b     E  L  T  T  R  I  L  A  T  K  K  K  K  S  Y  H  N  P  D  S  -
c      N  *  Q  Q  E  Y  *  Q  R  K  R  R  R  A  I  T  I  L  I  L -

     TTAAAGATTTCAAAAATTCCAGGTAAGAGAGATACATTCATTAAAATTCATATATTATAG
1081 ---------+---------+---------+---------+---------+---------+ 1140
     AATTTCTAAAGTTTTTAAGGTCCATTCTCTCTATGTAAGTAATTTTAAGTATATAATATC

a    L  K  I  S  K  I  P  G  K  R  D  T  F  I  K  I  H  I  L  *   -
b     *  R  F  Q  K  F  Q  V  R  E  I  H  S  L  K  F  I  Y  Y  S  -
c      K  D  F  K  N  S  R  *  E  R  Y  I  H  *  N  S  Y  I  I  V -

     TTTTTCATTTCACAGCTGTTATTTTCTTTTATCTTAACAATATTTTTTGATTAGCTGGAA
1141 ---------+---------+---------+---------+---------+---------+ 1200
     AAAAAGTAAAGTGTCGACAATAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT

a    F  F  I  S  Q  L  L  F  S  F  I  L  T  I  F  F  D  *  L  E   -
b     F  S  F  H  S  C  Y  F  L  L  S  *  Q  Y  F  L  I  S  W  K  -
c      F  H  F  T  A  V  I  F  F  Y  L  N  N  I  F  C  L  A  G  S -

     GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGGTAACTTAGCTTATTCACATTCAT
1201 ---------+---------+---------+---------+---------+---------+ 1260
     CATTTTTCATAGTTTATTCTCTTCGCGATCTGACTCCATTGAATCGAATAAGTGTAAGTA

a    V  K  S  I  K  *  E  K  R  *  T  E  V  T  *  L  I  H  I  H   -
b     *  K  V  S  N  K  R  S  A  R  L  R  *  L  S  L  F  T  F  I  -
c      K  K  Y  Q  I  R  E  A  L  D  C  G  N  L  A  Y  S  H  S  * -

     AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTTAAAAA
1261 ---------+---------+---------+---------+---------+---------+ 1320
     TCTAGCTGGAAGTATATAGGTTATGCTACTATTCCTTTGTCGTCAGTAGGCAAAATTTTT

a    R  S  T  F  I  Y  P  I  R  C  *  G  N  S  S  H  P  F  *  K   -
b     D  R  P  S  Y  I  Q  Y  D  D  K  E  T  A  V  I  R  F  K  N  -
c      I  D  L  H  I  S  N  T  M  I  R  K  Q  Q  S  S  V  L  K  I -

     TAGTGCTATGAGGACTAAATTTTTAGAGTCAAGAAATGGAGCCGAAATCTTAATCAAAAA
1321 ---------+---------+---------+---------+---------+---------+ 1380
     ATCACGATACTCCTGATTTAAAAATCTCAGTTCTTTACCTCGGCTTTAGAATTAGTTTTT

a    *  C  Y  E  D  *  I  F  R  V  K  K  W  S  R  N  L  N  Q  K   -
b     S  A  M  R  T  K  F  L  E  S  R  N  G  A  E  I  L  I  K  K  -
c      V  L  C  G  L  N  F  *  S  Q  E  M  E  P  K  S  *  S  K  R -
```

FIG. 35C

```
     GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA
1381 ---------+---------+---------+---------+---------+---------+ 1440
     CTTAACGCAGCTATAACGTTTTCTTAGCTTGAGATTTAGAAAGCAATTATTCATAATGGT

a   E  L  R  R  Y  C  K  R  I  E  L  *  I  F  R  *  *  V  L  P  -
b    N  C  V  D  I  A  K  E  S  N  S  K  S  F  V  N  K  Y  Y  Q  -
c     I  A  S  I  L  Q  K  N  R  T  L  N  L  S  L  I  S  I  T  N -

     ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA
1441 ---------+---------+---------+---------+---------+---------+ 1500
     TAGAACTAACTAACTTCTCTAACTGCTCCGTTGACGTGTCTTCTAGTAATTTCTTTATTT

a   I  L  I  D  C  R  D  C  R  G  N  C  T  E  D  H  *  R  N  K  -
b    S  C  L  I  E  E  I  D  E  A  T  A  Q  K  I  I  K  E  I  K  -
c     L  D  C  L  K  R  L  T  R  Q  L  H  R  R  S  L  K  K  *  S -

     GTAACTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTTCAA
1501 ---------+---------+---------+---------+---------+---------+ 1560
     CATTGAAAATAATTAATCTCTTATTTGATTTAATGATTATATCTCTAGTCGCTAGAAGTT

a   V  T  F  I  N  *  R  I  N  *  I  T  N  I  E  I  S  D  L  Q  -
b    *  L  L  L  I  R  E  *  T  K  L  L  I  *  R  S  A  I  F  N  -
c     N  F  Y  *  L  E  N  K  L  N  Y  *  Y  R  D  Q  R  S  S  I -

     TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT
1561 ---------+---------+---------+---------+---------+---------+ 1620
     AACTGCTTTATTTTCGACTTGATTTCAATCTGTTATTTTTTATGTTTGGAACCAGTTTTA

a   L  T  K  *  K  L  N  *  S  *  T  I  K  N  T  N  L  G  Q  N  -
b    C  R  N  K  S  C  T  K  V  R  Q  *  K  I  Q  T  L  V  K  I  -
c     D  E  I  K  A  E  L  K  L  D  N  K  K  Y  K  P  W  S  K  Y -

     ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAAATAAGGCAATAAATAAAATGA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     TAACTCCTTCCTTTTCTTCTGGTCAATCGTTTTCTTTTTTATTCCGTTATTTATTTTACT

a   I  E  E  G  K  E  D  Q  L  A  K  E  K  I  R  Q  *  I  K  C  -
b    L  R  K  E  K  K  T  S  *  Q  K  K  K  *  G  N  K  *  N  E  -
c     C  G  R  K  R  R  P  V  S  K  R  K  N  K  A  I  N  K  M  S -

     GTACAGAAGTGAAGAAATAAAAGATTTATTTTTTTCAATAATTTATTGAAAAGAGGGGTT
1681 ---------+---------+---------+---------+---------+---------+ 1740
     CATGTCTTCACTTCTTTATTTTCTAAATAAAAAAAGTTATTAAATAACTTTTCTCCCCAA

a   V  Q  K  C  R  N  K  R  F  I  F  F  N  N  L  L  K  R  G  V  -
b    Y  R  S  E  E  I  K  D  L  F  F  S  I  I  Y  C  K  E  G  F  -
c     T  E  V  K  K  *  K  I  Y  F  F  Q  *  F  I  E  K  R  G  F -

     TTGGGGTTTTGGGGTTTTGGGG
1741 ---------+---------+-- 1762
     AACCCCAAAACCCCAAAACCCC

a   L  G  F  W  G  F  G  -
b    W  G  F  G  V  L  G  -
c     G  V  L  G  F  W  -
```

FIG. 35D

```
  2  EVDVQNQADNHGIHSALKTCEEIKEAKTLYSWIQKVIRCRNQSQSHYKDL  51
     |:::  | :.:::|:  :|    |.:::|     ::. .  |  |..|.|.|
 19  ELELEMQENQNDIQVRVK....IDDPKQY..LVNVTAACLLQEGSYYQDK  62

 52  EDIKIFAQTNIVATPRDYNEEDFKVIARKEVF.STGLMIELIDKCLVELL 100
     ::      ..::    : .|.| ..|.. .|: ..:| |  ...::|.:
 63  DERRYIITKALL....EVAESDPEFICQLAVYIRNELYIRTTTNYIVAF. 107

101  SSSDVSDRQKLQCFGFQLKGNQLAKTHLLTALSTQKQYFFQDEWNQVRAM 150
                 ..: ....: ...: .|:  ..:.:    |:.||  :
108  .............CVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYI 144

151  IGNELFRHLYTKYLIFQRTSEGTLVQFCGNNVFDHLKVNDKFDKKQKGGA 200
     ::.. | :||              | .: :.:: ..|.... :::  ::
145  FDATEFKNLY.............LDRILSQDIRKELTFRKCLQRCVRSKF 181

201  ADMNE...PRCCSTCKYNVKNEKDHFLNNINVPNWNNMKSRTRIFYCTHF 247
     .::||   .::|..:  . |... ::|. .| ..|:. |.:
182  SEFNEYQLGKYCTES..QRKKTMFRYLSVTNKQKWDQTKKK......... 220

248  NRNNQFFKKHEFVSNKNNISAMDRAQTIFTNIFRFNRIRKKLKDKVIEKI 297
      |.: ...| :. .:.:: | .: :     :|:..:   | ||. |:.||
221  .RKENLLTKLQAIKESEDKSKRETG.....DIMNVEDAIKALKPAVMKKI 264

298  AYMLEKVKDFNFNYYLTKSCPLPENWRERKQKIENLINKTREEKSKYYEE 347
     | . . :|                    :. | ..: |.| |.|   :.:
265  AKRQNAMK..... .............KHMKAPKIPNSTLESKYLTFKD 294

348  LFSYTTDNKCVTQFINEFFYNILPKDFLTGRNRKNFQKKVKKYVELNKHE 397
     |:..   ...     .| .|.||.|.: .. :  .  .  ..:| .|
295  LIKFCHISEP.....KERVYKILGKKYPKTEEEYKAAFGDSASAPFN.PE 338

398  LIHKNLLLEKINTREISWMQVETSAKHFYYFDHENIYVLWKLLRWIFEDL 447
     |  |.. :|  .|:| .: . :..|. :  :  .| ...  :|| : :
339  LAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSN.. 386

448  VVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSIADLKKETLAEVQE 497
                                    |:| ::.|
387  ..............................ILKAGVSD.......... 394

498  KEVEEWKKSLGFAPGKLRLIPKKTTFRPIMTFNKKIVNSDRKTTKLTTNT 547
                                                    ||:.
395  ..............................................TTHS 398

548  KLLNSHLMLKTLKNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL 597
        |                                  :|. |.|:..|:
399  IVINK..................................ICEPKAVENSKM 415

598  FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKNNIVID 647
     |    :. .. :.|| | :.. :|..|    .:   :.:||    |:   |::
416  F..PLQFFSAIEAVN.EAVTKGFKAKK...RENMNLKGQIEAVKE..VVE 457

648  SKNFRKKEMKDYFRQKFQKIALEGGQYPTLFSVLENEQNDLNAKKTLIVE 697
     ..:  ||:|            .||..: ..:..| |. ..:. .. .| :.
458  KTDEEKKDM...........ELEQTEEGEFVKVNEGIGKQYINSIELAIK 496

698  AKQRNYFKKDNLLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSS 747
        .. :.. .   :::. .  .  .  :   |.  :::. .| .. :|:
497  IAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLECALVLGL 546

748  FYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLITTQENNAVLFI 797
     :  .  |.||: :::...| ..:.. ::: :
547  MVKQRCEKSSFYIFSSPSSQCNKCYLEVDL..................... 576
```

FIG. 36A

```
798  EKLINVSRENGFKFNMKK.LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846
577  .......PGDELRPSMQKLLQEKGKLGGG..TDFPYECIDEWTKNKTHVD 617

847  WIGISIDMKTLALMPNINLRIEGILCTLNLNMQTKKASMWLKKKLKSFLM 896
618  NIVILSDMMIAEGYSDINVRGSSIVNSI.............KKYKDEVN 653

897  NNITHYFRKTITTEDFANKTLNKLFISGGYKYMQCAKEYKD.HFKKNLAM 945
654  PNIKIF...AVDLEGYG.............KCLNLGDEFNENNYIKIFGM 687

946  SSMIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFIE 995
688  SDSI...............LKFISAKQGGA...............NMVE 706

996  IFSTKKYIFNRVC 1008
707  VI..KNFALQKIG 717
```

FIG. 36B

```
132  LSTQKQYFFQDEWNQVRAMIGNEL.FRHLYTKYLIFQRTSE..GTLVQFC 178
  1  MSRRNQ.......KKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQQI  43

179  GNNVFDHLKVNDKFDKKQKGGAADMNEPRCCSTCKYNVKNEKDHFLNNIN 228
 44  KEEDLKLLKFKNQDQDGNSGNDDDDEE.........NNSNKQQELLRRVN  84

229  VPNWNNMKSRTRIFYCTHFNRNNQFFKKHEFVSNKNNISAMDRAQTIFTN 278
 85  .................QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN 114

279  IFRFNRIRKKLKDKVIEKIAYMLEKVKDFNFNYYLTKSCPLPENWRERKQ 328
115  GLSEQQVKEEQLRTITEEQVKYQNLVFNMDYQLDLNESGGHRRHRRETDY 164

329  KIENLINKTREEKSKYYEELFSYTTDNKCVTQFINE.FFYNILPKDFLTG 377
165  DTEKWFEISHDQK.............NYVSIYANQKTSYCWWLKDYFNK  200

378  RNRKNFQKKVKKYVELNKHELIHKNLLLEKINTREISWMQVETSAKHFYY 427
201  NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNSYQTVNID......  242

428  FDHENIYVLWKLLRWI..FEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNI 475
243  VNFDNNLCILALLRFLLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290

476  WDVIMKMSIADLKKETLAEVQEKEVEEWKKSLGFAPGKLRLIPKKTTFRP 525
291  FAVVFSHR.........HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ  330

526  IMTFNKKIVNSDRKTTKLTTNTKLLNSHLMLKTLKNRMFKDPFGFAVFNY 575
331  VYSFSTDLKLVD..TNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENL 378
```

FIG. 37A

```
576  DDVMKKYEEFVCKWKQVGQPKLF. . ....FATMDIEKCYDS..VNREK  615
379  NVLLKKVKH  ANLNLVSIPTQFNFDFYFVNLQHLKLEFGLEPNILTKQK  426

516  LSTFL......KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKKEMK  657
427  LENLLLSIKQSKNLKFLRLNFYTYVAQETSRKQILKQATTIKNLKNNKNQ  476

558  DYFRQKFQKIALEGGQYPTLFSVLEN..EQNDLNAKKTLIVEAKQRNYFK  705
477  EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY..  520

706  KDNLLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSSFYYATLEE  755
521  .DSLHKLLIRSTNLKKFKLSYKYEMEKSKMDTFIDLKNI.....YETLNN  564

756  SSLGFLRDESMNPENPNVNLLMRLTDDYLLITTQENNAVLFIEKLINVSR  305
565  .....LKRCSVNISNPHGNISYELTN.. .....KDSTFYKFKLTLNQE  500

806  ENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK  855
601  LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDSLCKSIASCKNLQ  648

856  TLALMPNINLRIEGILCTLNLNMQT..KKASMWLKK..KLKSFLMNNITH  901
649  NVNI.......IASLLYPNNIQKNPFNKPNLLFFKQFEQLKNLENVSINC  691

902  YFRKTI...TTEDFANKTLNKLFISGGYKYMQCAKEYKDHFKKNLAMSSM  948
692  ILDQHILNSISEFLEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPEL  741

949  IDLEVSKIIYSVT..............RAFFKYLVCNIKDT..IFGEEHY  982
742  NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDF  791

983  PDFFLS  TLKHFIEIFSTKKY IFNRVCMILKAKEAKLKSDQCQSLIQ  1028
792  DQNTVSDDSIKKILESISESKYHHYLRLNPSQSSSLIKSENEEIQELLK  840
```

FIG. 37B

```
  4  DIDLDDIENLLPNTFNKYSSSCSDKKGCKTLKSGSKSPSLTIPK......  47
617  NVKSAKIESSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNP  666

 48  .......LQKQLEFYFSDANLYNDSFLRKLVLKSGEQRVE....IETLLM  86
667  FNKPNLLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL  716
```

FIG. 38

```
  1  MEMDIDLDDIENL.....LPNTFNKYSSSCSDKKGCKTLKSGSKSPS...   42
491  IELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLEC  540

 43  .LTIPKLQKQ......LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL   85
541  ALVLGLMVKQRCEKSSFYIFSSPSSQCNKCYL.EVDLPGDELRPSMQKLL  589
```

FIG. 39

```
telomerase p43   LQKQLEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLLM
human La         ICHQUEYYFGDFNLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA      ICEQIEYYFGDHNLPRDKFLKQQI.LLDDGWVPLETMIK
Drosophila La    ILRQVEYYFGDANLNRDKFLREQIGKNEDGWVPLSVLVT
S. c. Lhp1p      CLKQVEFYFSEFNFPYDRFLRTTAEK.NDGWVPISTIAT
```

FIG. 41

```
   1 aactcattta attactaatt taatcaacaa gattgataaa aagcagtaaa taaaacccaa
  61 tagatttaat ttagaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
 121 gccaaaagcc gaaaaattgt ggtgggaact tgaattagag atgcaagaaa accaaaatga
 181 tatataagtt agggttaaga ttgacgatcc taagcaatat ctcgtgaacg tcactgcagc
 241 atgtttgttg taggaaggta gttactacta agataaagat gaaagaagat atatcatcac
 301 taaagcactt cttgaggtgg ctgagtctga tcctgagttc atctgctagt tggcagtcta
 361 catccgtaat gaactttaca tcagaactac cactaactac attgtagcat tttgtgttgt
 421 ccacaagaat actcaaccat tcatcgaaaa gtacttcaac aaagcagtac ttttgcctaa
 481 tgacttactg gaagtctgtg aatttgcata ggttctctat atttttgatg caactgaatt
 541 caaaaatttg tatcttgata ggatactttc ataagatatt cgtaaggaac tcactttccg
 601 taagtgttta caaagatgcg tcagaagcaa gttttctgaa ttcaacgaat actaacttgg
 661 taagtattgc actgaatcct aacgtaagaa aacaatgttc cgttacctct cagttaccaa
 721 caagtaaaag tgggattaaa ctaagaagaa gagaaaagag aatctcttaa ccaaacttta
 781 ggcaataaag gaatctgaag ataagtccaa gagagaaact ggagacataa tgaacgttga
 841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaa atagccaaga gatagaatgc
 901 catgaagaaa cacatgaagg cacctaaaat tcctaactct accttggaat caaagtactt
 961 gaccttcaag gatctcatta agttctgcca tatttctgag cctaaagaaa gagtctataa
1021 gatccttggt aaaaaatacc ctaagaccga agaggaatac aaagcagcct ttggtgattc
1081 tgcatctgca cccttcaatc ctgaattggc tggaaagcgt atgaagattg aaatctctaa
1141 aacatgggaa aatgaactca gtgcaaaagg caacactgct gaggtttggg ataatttaat
1201 ttcaagcaat taactcccat atatggccat gttacgtaac ttgtctaaca tcttaaaagc
1261 cggtgtttca gatactacac actctattgt gatcaacaag atttgtgagc ccaaggccgt
1321 tgagaactcc aagatgttcc ctcttcaatt ctttagtgcc attgaagctg ttaatgaagc
1381 agttactaag ggattcaagg ccaagaagag agaaaatatg aatcttaaag gtcaaatcga
1441 agcagtaaag gaagttgttg aaaaaaccga tgaagagaag aaagatatgg agttggagta
1501 aaccgaagaa ggagaatttg ttaaagtcaa cgaaggaatt ggcaagcaat acattaactc
1561 cattgaactt gcaatcaaga tagcagttaa caagaattta gatgaaatca aaggacacac
1621 tgcaatcttc tctgatgttt ctggttctat gagtacctca atgtcaggtg gagccaagaa
1681 gtatggttcc gttcgtactt gtctcgagtg tgcattagtc cttggtttga tggtaaaata
1741 acgttgtgaa aagtcctcat tctacatctt cagttcacct agttctcaat gcaataagtg
1801 ttacttagaa gttgatctcc ctggagacga actccgtcct tctatgtaaa aacttttgca
1861 agagaaagga aaacttggtg gtggtactga tttcccctat gagtgcattg atgaatggac
1921 aaagaataaa actcacgtag acaatatcgt tattttgtct gatatgatga ttgcagaagg
1981 atattcagat atcaatgtta gaggcagttc cattgttaac agcatcaaaa agtacaagga
2041 tgaagtaaat cctaacatta aaatctttgc agttgactta gaaggttacg gaaagtgcct
2101 taatctaggt gatgagttca atgaaaacaa ctacatcaag atattcggta tgagcgattc
2161 aatcttaaag ttcatttcag ccaagcaagg aggagcaaat atggtcgaag ttatcaaaaa
2221 ctttgccctt caaaaaatag gacaaaagtg agtttcttga gattcttcta taacaaaaat
2281 ctcaccccac ttttttgttt tattgcatag ccattatgaa atttaaatta ttatctattt
2341 atttaagtta cttacatagt ttatgtatcg cagtctatta gcctattcaa atgattctgc
2401 aaagaacaaa aaagattaaa a
```

FIG. 42

```
                          Motif A                          Motif B

Consensus                   h--hDh---h--h                    h---+-QG---SP
telomerase p123   GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL-100-KFYKQTKGIPQGLCVSSILSSFYYATLEESSLGFL
Dong (LINE)       KNRNLHCTYDDYKKAFDSIPHSWLIQVLEIYKIN- 28-RQIAIKKGIYQGDSLSPLWFCLALNPLSHQLHNDR
a1 S.c. (groupII) FGGSNWFREVDLKKCFDTISHDLIIKELKRYISD- 26-HVPVGPRVCVQGAPTSPALCNAVLLRLDRRLAGLA
HIV-RT            LKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP-  7-GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQN
L8543.12          VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN- 68-KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK

                          Motif C                          Motif D          Motif E

Consensus                 h--YhDDhhh                       Gh-h---K         h-hLGh-h
telomerase p123   -14-LMRLTDDYLLITTQENN-0-AVLFIEKLINVSRENGFKFNMKKLQT-23-QDYCDWIGISI
Dong (LINE)       -16-HLIYMDDIKLYAKNDKE-0-MKKLIDTTTIFSNDISMQFGLDKCKT-25-KCLYKYLGFQQ
a1 S.c. (groupII) -55-YVRYADDILIGVLGSKN-2-KIIKRDLNNFLNS.LGLTINEEKTLI- 4-ETPARFLGYNI
HIV-RT            - 4-IYQYMDDLYVGSHLEIG-1-HRTKIEELRQHLLRWGLTTPDKKHQK- 0-EPPFLWMGYEL
L8543.12          - 8-ILKLADDFLIISTDQQQ.........VINIKKLAMGGFQKYNAKANR-41-IRSKSSKGIFR
```

FIG. 40

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAACLLQEGSYYQDKDERRYIITKALLEVAESDPEFICQLA
VYIRNELYIRTTTNYIVAFCVVHKNTQPFIEKYFNKAVLLPNDL
LEVCEFAQVLYIFDATEFKNLYLDRILSQDIRKELTFRKCLQRC
VRSKFSEFNEYQLGKYCTESQRKKTMFRYLSVTNKQKWDQTKKK
RKENLLTKLQAIKESEDKSKRETGDIMNVEDAIKALKPAVMKKI
AKRQNAMKKHMKAPKIPNSTLESKYLTFKDLIKFCHISEPKERV
YKILGKKYPKTEEEYKAAFGDSASAPFNPELAGKRMKIEISKTW
ENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGVSDTT
HSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKR
ENMNLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIG
KQYINSIELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGA
KKYGSVRTCLECALVLGLMVKQRCEKSSFYIFSSPSSQCNKCYL
EVDLPGDELRPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHV
DNIVILSDMMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFA
VDLEGYGKCLNLGDEFNENNYIKIFGMSDSILKFISAKQGGANM
VEVIKNFALQKIGQK

FIG. 43

MSRRNQKKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQQIK
EEDLKLLKFKNQDQDGNSGNDDDDEENNSNKQQELLRRVNQIKQ
QVQLIKKVGSKVEKDLNLNEDENKKNGLSEQQVKEEQLRTITEE
QVKYQNLVFNMDYQLDLNESGGHRRHRRETDYDTEKWFEISHDQ
KNYVSIYANQKTSYCWWLKDYFNKNNYDHLNVSINRLETEAEFY
AFDDFSQTIKLTNNSYQTVNIDVNFDNNLCILALLRFLLSLERF
NILNIRSSYTRNQYNFEKIGELLETIFAVVFSHRHLQGIHLQVP
CEAFQYLVNSSSQISVKDSQLQVYSFSTDLKLVDTNKVQDYFKF
LQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNLVSIP
TQFNFDFYFVNLQHLKLEFGLEPNILTKQKLENLLLSIKQSKNL
KFLRLNFYTYVAQETSRKQILKQATTIKNLKNNKNQEETPETKD
ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDSLHKLLI
RSTNLKKFKLSYKYEMEKSKMDTFIDLKNIYETLNNLKRCSVNI
SNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFN
NVKSAKIESSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPN
NIQKNPFNKPNLLFFKQFEQLKNLENVSINCILDQHILNSISEF
LEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPELNQVYINQ
QLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDFD
QNTVSDDSIKKILESISESKYHHYLRLNPSQSSSLIKSENEEIQ
ELLKACDEKGVLVKAYYKFPLCLPTGTYYDYNSDRW

FIG. 45

MKILFEFIQDKLDIDLQTNSTYKENLKCGHFNGLDEILTTCFAL
PNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKI
ARNEDVNNSLFCHSANVNVTLLKGAAWKMFHSLVGTYAFVDLLI
NYTVIQFNGQFFTQIVGNRCNEPHLPPKWVQRSSSSSATAAQIK
QLTEPVTNKQFLHKLNINSSSFFPYSKILPSSSSIKKLTDLREA
IFPTNLVKIPQRLKVRINLTLQKLLKRHKRLNYVSILNSICPPL
EGTVLDLSHLSRQSPKERVLKFIIVILQKLLPQEMFGSKKNKGK
IIKNLNLLLSLPLNGYLPFDSLLKKLRLKDFRWLFISDIWFTKH
NFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVTIVYFR
HDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSKMRI
IPKKSNNEFRIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEY
LRNKRPTSFTKIYSPTQIADRIKEFKQRLLKKFNNVLPELYFMK
FDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFNTNTGVL
KLFNVVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTA
LWVEDKCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS
QDTLILKLADDFLIISTDQQQVINIKKLAMGGFQKYNAKANRDK
ILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMNNFHIRSKS
SKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISE
CYKSAFKDLSINVTQNMQFHSFLQRIIEMTVSGCPITKCDPLIE
YEVRFTILNGFLESLSSNTSKFKDNIILLRKEIQHLQAYIYIYI
HIVN

FIG. 46

```
   1 tcaatactat taattaataa ataaaaaaaa gcaaactaca aagaaaatgt caaggcgtaa
  61 ctaaaaaaag ccataggctc ctataggcaa tgaaacaaat cttgattttg tattacaaaa
 121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
 181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
 241 tgatgatgat gaagaaaaca actcaaataa ataataagaa ttattaagga gagtcaatta
 301 gattaagtag caagtttaat tgataaaaaa agttggttct aaggtagaga aagatttgaa
 361 tttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta
 421 attaagaacg attactgaag aataggttaa gtattaaaat ttagtattta acatggacta
 481 ccagttagat ttaaatgaga gtggtggcca tagaagacac agaagagaaa cagattatga
 541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa tttacgccaa
 601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaca attatgatca
 661 tcttaatgta agcattaaca gactagaaac tgaagccgaa ttctatgcct ttgatgattt
 721 ttcacaaaca atcaaactta ctaataattc ttactagact gttaacatag acgttaattt
 781 tgataataat ctctgtatac tcgcattgct tagattttta ttatcactag aaagattcaa
 841 tattttgaat ataagatctt cttatacaag aaattaatat aattttgaga aaattggtga
 901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cacttacaag gcattcattt
 961 acaagttcct tgcgaagcgt tctaatattt agttaactcc tcatcataaa ttagcgttaa
1021 agatagctaa ttataggtat actctttctc tacagactta aaattagttg acactaacaa
1081 agtccaagat tattttaagt tcttataaga attccctcgt ttgactcatg taagctagta
1141 ggctatccca gttagtgcta ctaacgctgt agagaacctc aatgttttac ttaaaaaggt
1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctactt
1261 tgttaattta taacatttga aattagagtt tggattagaa ccaaatattt tgacaaaaca
1321 aaagcttgaa aatctacttt tgagtataaa ataatcaaaa aatcttaaat ttttaagatt
1381 aaacttttac acctacgttg cttaagaaac ctccagaaaa cagatattaa aacaagctac
1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaaagatga
1501 aactccaagc gaaagcacaa gtggtatgaa atttttgat catctttctg aattaaccga
1561 gcttgaagat ttcagcgtta acttgtaagc tacccaagaa atttatgata gcttgcacaa
1621 acttttgatt agatcaacaa atttaaagaa gttcaaatta agttacaaat atgaaatgga
1681 aaagagtaaa atggatacat tcatagatct taagaatatt tatgaaacct taaacaatct
1741 taaaagatgc tctgttaata tatcaaatcc tcatggaaac atttcttatg aactgacaaa
1801 taaagattct acttttata aatttaagct gaccttaaac taagaattat aacacgctaa
1861 gtatactttt aagtagaacg aatttttaatt taataacgtt aaaagtgcaa aaattgaatc
1921 ttcctcatta gaaagcttag aagatattga tagtctttgc aaatctattg cttcttgtaa
1981 aaatttacaa aatgttaata ttatcgccag tttgctctat cccaacaata tttagaaaaa
2041 tcctttcaat aagcccaatc ttctattttt caagcaattt gaataattga aaaatttgga
2101 aaatgtatct atcaactgta ttcttgatca gcatatactt aatttctattt cagaattctt
2161 agaaaagaat aaaaaaataa aagcattcat tttgaaaaga tattatttat tacaatatta
2221 tcttgattat actaaattat ttaaaacact tcaatagtta cctgaattaa attaagttta
2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaaa
2341 ccacaagcaa aaagctttct atgaaccatt atgtgagttt atcaaagaat catcctaaac
2401 cctttagcta atagattttg accaaaacac tgtaagtgat gactctatta aaaagatttt
2461 agaatctata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttaatctag
2521 cagtttaatt aaatctgaaa acgaagaaat ttaagaactt ctcaaagctt gcgacgaaaa
2581 aggtgtttta gtaaaagcat actataaatt ccctctatgt ttaccaactg gtacttatta
2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaaataaa tattaaatat
2701 tgaatatttc tttgcttatt atttgaataa tacatacaat agtcattttt agtgttttga
2761 atatatttta gttatttaat tcattatttt aagtaaataa ttatttttca atcatttttt
2821 aaaaaatcg
```

FIG. 44

Oxytricha	LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Euplotes	LCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIG. 47

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTTCGCACTACCAAATTCAAGAAAAATAGCATTACCATGCCTTCCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT
ATACAACAACGTACTAACATTTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG
TCTTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGGAAAAT
GTTCCACAGTTTGGTCGGTACATACGCATTCGTTGATTTATTGATCAATTATACAGTAAT
TCAATTTAATGGGCAGTTTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT
GCCGCCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA
ACTTACAGAACCAGTGACAAATAAACAATTCTTACACAAGCTCAATATAAATTCCTCTTC
TTTTTTTCCTTATAGCAAGATCCTTCCTTCATCATCATCTATCAAAAAGCTAACTGACTT
GAGAGAAGCTATTTTTCCCACAAATTTGGTTAAAATTCCTCAGAGACTAAAGGTACGAAT
TAATTTGACGCTGCAAAAGCTATTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT
GAATAGTATTTGCCCACCATTGGAAGGGACCGTATTGGACTTGTCGCATTTGAGTAGGCA
ATCACCAAAGGAACGAGTCTTGAAATTTATCATTGTTATTTTACAGAAGTTATTACCCCA
AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATTATCAAGAATCTAAATCTTTTATT
AAGTTTACCCTTAAATGGCTATTTACCATTTGATAGTTTGTTGAAAAAGTTAAGATTAAA
GGATTTTCGGTGGTTGTTCATTTCTGATATTTGGTTCACCAAGCACAATTTTGAAAACTT
GAATCAATTGGCGATTTGTTTCATTTCCTGGCTATTTAGACAACTAATTCCCAAAATTAT
ACAGACTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA
TGATACTTGGAATAAACTTATCACCCCTTTTATCGTAGAATATTTTAAGACGTACTTAGT
CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA
AATGAGGATTATACCAAAAAAAAGTAATAATGAGTTCAGGATTATTGCCATCCCATGCAG
AGGGGCAGACGAAGAAGAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC
CACTCAAAAAATTTTAGAATACCTAAGAAACAAAAGGCCGACTAGTTTTACTAAAATATA
TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTTAAGCAGAGACTTTTAAAGAAATT
TAATAATGTCTTACCAGAGCTTTATTTCATGAAATTTGATGTCAAATCTTGCTATGATTC
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAAATGGGTT
TTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTAAAACAGCTTT
GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTCAGGGCTCTAGTTTATC
TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTTATAGCGAGTTTAAAGC
CAGTCCTAGCCAGGACACATTAATTTTAAAACTGGCTGACGATTTCCTTATAATATCAAC
AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAATATAA
TGCGAAAGCCAATAGAGACAAAATTTTAGCCGTAAGCTCCCAATCAGATGATGATACGGT
TATTCAATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAAACATTCAAG
CACAATGAATAATTTCCATATCCGTTCGAAATCTAGTAAAGGGATATTTCGAAGTTTAAT
AGCGCTGTTTAACACTAGAATCTCTTATAAAACAATTGACACAAATTTAAATTCAACAAA
CACCGTTCTCATGCAAATTGATCATGTTGTAAAGAACATTTCGGAATGTTATAAATCTGC
TTTTAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTTCATTCGTTCTTACAACG
CATCATTGAAATGACAGTCAGCGGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA
TGAGGTACGATTCACCATATTGAATGGATTTTTGGAAAGCCTATCTTCAAACACATCAAA
ATTTAAAGATAATATCATTCTTTTGAGAAAGGAAATTCAACACTTGCAAGC

FIG. 48

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREK
RAERLTSRVKALFSVLNYERA

FIG. 49

GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTC
TTTCTTTTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAG
AGGGTGCAGCTGCGGGACGTGTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGC
CAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGC
TGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTTCCGCAGA
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGCGTGCT
CAACTACGAGCGGGCGCG

FIG. 50

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
DSKPDEGVQFSSPKCSQSELIANVVKQMFDESFERRRNLLMKGFSMNHEDFRAMHVNGVQNDLVSTF
PNYLISILESKNWQLLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKKRKR
TIETSITQNKSARKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG
LINAFQVKQLHKVIPLVSQSTVVPKRLLKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS
YSLKPNQVFAFLRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWL
VLGKRSNAKMCLSDFEKRKQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCR
PFITSMKMEAFEKINENNVRMDTQKTTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTN
QTLRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLM
FRIVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVVQLLSMKTSDTLFVDFVDY
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENSNGIINNTFFNESKKRMPFFG
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN
IYRLGYSMCMRAQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFC
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFLHRRIAD

FIG. 51

ggtaccgatttactttcctttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaaga
actcaataacaataccaagtcaaattccaatatgaaggtgttattagtgatcgataatatttctatttttatcggtcgtta
ccaagtataaggacaaaaagaacaacttccttcccctaaagacttttacttttattaatttacttttcaaatatatttcg
ggttcgcttacttttaatcgtggtactgttttagctgctacttctagccaaccgcgtgtttctaccccgtcattggatat
agctcttggagtagctcacagaaatccttacaaatcttctgatgagactatattagattcattacagtccgtgcatattc
ttaacatggagccttacactttagatgagtcacgtcgcatgatggagtatttggtatcatccaacgtttgccttgaaaag
gttgataattatttgcaaaatcatgtccttagtggtggtaatccgcgaaagtttttgatgcttgcacacgtctagcatg
attgagatattcaaaaatttctatccactacaactcctttaacgcggttttatttttctattttctattctcatgttgtt
ccaaatatgtatcatctcgtattaggctttttccgttttactcctggaatcgtacctttttcactattccccctaatga
ataatctaaattagtttcgcttataattgatagtagtagaaagattggtgattctactcgtgtaatgttattagtttaaa
gatactttgcaaaacatttattagctatcattatataaaaaaaatcctataattataaatattaatcaatatttgcggtc
actatttatttaaaacgttatgatcagtaggacactttgcatatatatagttatgcttaatggttacttgtaacttgcAT
GACCGAACACCATACCCCCAAAAGCAGGATTCTTCGCTTTCTAGAGAATCAATATGTATACCTATGTACCTTAAATGATT
ATGTACAACTTGTTTTGAGAGGGTCGCCGGCAAGCTCGTATAGCAATATATGCGAACGCTTGAGAAGCGATGTACAAACG
TCCTTTTCTATTTTTCTTCATTCGACTGTAGTCGGCTTCGACAGTAAGCCAGATGAAGGTGTTCAATTTTCTTCTCCAAA
ATGCTCACAGTCAGAGgtatatatatttttgttttgatttttttctattcgggatagctaatatatgggcagCTAATAGC
GAATGTTGTAAAACAGATGTTCGATGAAAGTTTTGAGCGTCGAAGGAATCTACTGATGAAAGGGTTTTCCATGgtaaggt
attctaattgtgaaatatttacctgcaattactgtttcaaagagattgtatttaaccgataaagAATCATGAAGATTTTC
GAGCCATGCATGTAAACGGAGTACAAAATGATCTCGTTTCTACTTTTCCTAATTACCTTATATCTATACTTGAGTCAAAA
AATTGGCAACTTTTGTTAGAAATgtaaataccggttaagatgttgcgcactttgaacaagactgacaagtatagTATCGG
CAGTGATGCCATGCATTACTTATTATCCAAAGGAAGTATTTTTGAGGCTCTTCCAAATGACAATTACCTTCAGATTTCTG
GCATACCACTTTTTAAAAATAATGTGTTTGAGGAAACTGTGTCAAAAAAAAAGAAAGCGAACCATTGAAACATCCATTACT
CAAAATAAAAGCGCCCGCAAAGAAGTTTCCTGGAATAGCATTTCAATTAGTAGGTTTAGCATTTTTTACAGGTCATCCTA
TAAGAAGTTTAAGCAAGgtaactaatactgttatccttcataactaattttagATCTATATTTTAACTTACACTCTATTT
GTGATCGGAACACAGTACACATGTGGCTTCAATGGATTTTTCCAAGGCAATTTGGACTTATAAACGCATTTCAAGTGAAG
CAATTGCACAAAGTGATTCCACTGGTATCACAGAGTACAGTTGTGCCCAAACGTCTCCTAAAGGTATACCCTTTAATTGA
ACAAACAGCAAAGCGACTCCATCGTATTTCTCTATCAAAAGTTTACAACCATTATTGCCCATATATTGACACCCACGATG
ATGAAAAAATCCTTAGTTATTCCTTAAAGCCGAACCAGGTGTTTGCGTTTCTTCGATCCATTCTTGTTCGAGTGTTTCCT
AAATTAATCTGGGGTAACCAAAGGATATTTGAGATAATATTAAAAGgtattgtataaaatttattaccactaacgatttt
accagACCTCGAAACTTTCTTGAAATTATCGAGATACGAGTCTTTTAGTTTACATTATTTAATGAGTAACATAAAGgtaa
tatgccaaatttttttaccattaattaacaatcagATTTCAGAAATTGAATGGCTAGTCCTTGGAAAAAGGTCAAATGCG
AAAATGTGCTTAAGTGATTTTGAGAAACGCAAGCAAATATTTGCGGAATTCATCTACTGGCTATACAATTCGTTTATAAT
ACCTATTTTACAATCTTTTTTTTTATATCACTGAATCAAGTGATTTACGAAATCGAACTGTTTATTTTAGAAAAGATATTT
GGAAACTCTTGTGCCGACCCTTTATTACATCAATGAAAATGGAAGCGTTTGAAAAAATAAACGAGgtattttaaagtatt
ttttgcaaaaagctaatattttcagAACAATGTTAGGATGGATACTCAGAAAACTACTTTGCCTCCAGCAGTTATTCGTC
TATTACCTAAGAAGAATACCTTTCGTCTCATTACGAATTTAAGAAAAAGATTCTTAATAAAGgtattaatttttggtcat
caatgtactttacttctaatctattattagcagATGGGTTCAAACAAAAAAATGTTAGTCAGTACGAACCAAACTTTACG
ACCTGTGGCATCGATACTGAAACATTTAATCAATGAAGAAAGTAGTGGTATTCCATTTAACTTGGAGGTTTACATGAAGC

FIG. 52A

TTCTTACTTTTAAGAAGGATCTTCTTAAGCACCGAATGTTTGGgtaattatataatgcgcgattcctcattattaatttt
gcagGCGTAAGAAGTATTTTGTACGGATAGATATAAAATCCTGTTATGATCGAATAAAGCAAGATTTGATGTTTCGGATT
GTTAAAAAGAAACTCAAGGATCCCGAATTTGTAATTCGAAAGTATGCAACCATACATGCAACAAGTGACCGAGCTACAAA
AAACTTTGTTAGTGAGGCGTTTTCCTATTgtaagtttatttttcattggaatttttaacaaattctttttagTTGAT
ATGGTGCCTTTTGAAAAAGTCGTGCAGTTACTTTCTATGAAAACATCAGATACTTTGTTTGTTGATTTTGTGGATTATTG
GACCAAAAGTTCTTCTGAAATTTTTAAAATGCTCAAGGAACATCTCTCTGGACACATTGTTAAGgtataccaattgttga
attgtaataacactaatgaaactagATAGGAAATTCTCAATACCTTCAAAAAGTTGGTATCCCTCAGGGCTCAATTCTGT
CATCTTTTTTGTGTCATTTCTATATGGAAGATTTGATTGATGAATACCTATCGTTTACGAAAAAGAAAGGATCAGTGTTG
TTACGAGTAGTCGACGATTTCCTCTTTATAACAGTTAATAAAAAGGATGCAAAAAAATTTTTGAATTTATCTTTAAGAGg
tgagttgctgtcattcctaagttctaaccgttgaagGATTTGAGAAACACAATTTTTCTACGAGCCTGGAGAAAACAGTA
ATAAACTTTGAAAATAGTAATGGGATAATAAACAATACTTTTTTTAATGAAAGCAAGAAAAGAATGCCATTCTTCGGTTT
CTCTGTGAACATGAGGTCTCTTGATACATTGTTAGCATGTCCTAAAATTGATGAAGCCTTATTTAACTCTACATCTGTAG
AGCTGACGAAACATATGGGGAAATCTTTTTTTTACAAAATTCTAAGgtatactgtgtaactgaataatagctgacaaata
atcagATCGAGCCTTGCATCCTTTGCACAAGTATTTATTGACATTACCCACAATTCAAAATTCAATTCTTGCTGCAATAT
ATATAGGCTAGGATACTCTATGTGTATGAGAGCACAAGCATACTTAAAAAGGATGAAGGATATATTTATTCCCCAAAGAA
TGTTCATAACGGgtgagtacttattttaactagaaaagtcattaattaaccttagATCTTTTGAATGTTATTGGAAGAAA
AATTTGGAAAAAGTTGGCCGAAATATTAGGATATACGAGTAGGCGTTTCTTGTCCTCTGCAGAAGTCAAATGgtacgtgt
cggtctcgagacttcagcaatattgacacatcagGCTTTTTTGTCTTGGAATGAGAGATGGTTTGAAACCCTCTTTCAAA
TATCATCCATGCTTCGAACAGCTAATATACCAATTTCAGTCATTGACTGATCTTATCAAGCCGCTAAGACCAGTTTTGCG
ACAGGTGTTATTTTTACATAGAAGAATAGCTGATTAAtgtcattttcaatttattatatacatcctttattactggtgtc
ttaaacaatattattactaagtatagctgacccccaaagcaagcatactataggatttctagtaaagtaaaattaatctc
gttattagttttgattgacttgtctttatccttatacttttaagaaagattgacagtggttgctgactactgcccacatg
cccattaaacgggagtggttaaacattaaaagtaatacatgaggctaatctcctttcatttagaataaggaaagtggttt
tctataatgaataatgcccgcactaatgcaaaaagacgaagattatcttctaaacaaggggattaagcatatccgaagg
aaaagagagtaatatacccagtgttgttgaagaaagcaaggataatttggaacaagcttctgcagatgacaggctaaatt
ttggtgaccgaattttggtaaaagccccaggttatccatggtggccggccttgctactgagacgaaaagaaactaaggat
agtttgaatactaatagctcatttaatgtcttatataaggttttgttttttcctgacttcaattttgcatgggtgaaaag
aaatagtgttaagccattattggattccgaaatagccaaatttcttggttcctcaaagcggaagtctaaagaacttattg
aagcttatgaggcttcaaaaactcctcctgatttaaaggaggaatcttccaccgatgaggaaatggatagcttatcagct
gctgaggagaagcctaatttttgcaaaaaagaaaatatcattgggagacatctcttgatgaatcagatgcggagagtat
ctccagcggatccttgatgtcaataacttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctcta
cgcagttaagtgaccaaaggtacc

FIG. 52B

```
EST2 pep                FFYCTEISST VTIVYFRHDT WN----KLIT P-----FIVE YFK-TYLVEN    40
Euplotes pep            FFYVTEQQKS YSKTYYYRKN IWDVI-MKMS IAD----LKK ETLA--EVQE    43
Trans of tetrahymen     ------KHKE GSQIFYYRKP IWKLVSKLTI VKVRIQFSEK NKQMKNNFYQ    44

Consensus               FFY.TE..K. .S..YYYRK. IW...-KL.. ...---F..K .......V..    50


EST2 pep                NVCRNHNSY- ---------- TLSNFNHSKM RIIPKKSNNE FRIIAIPCRG    79
Euplotes pep            KEVEEWKKSL ---------- ---GFAPGKG RLIPKKTT-- FRPIMTFNKK    78
Trans of tetrahymen     KIQLEEENLE KVEEKLIPED SFQKYPQGKL RIIPKKGS-- FRPIMTFLRK    92

Consensus               K...E..... ---------- ....F..GKL RIIPKK..-- FRPIMTF.RK   100


EST2 pep                ADEEEFTIYK ENHKNAIQPT QKILEYLRNK RPTSFTKIYS PTQIADRIKE   129
Euplotes pep            IVNSDRKTTK LTTNTKLLNS HLMLKTLKN- ------RMFK -DPFGFAVFN   120
Trans of tetrahymen     DKQKNIK--- LNLNQILMDS QLVFRNLKD- ------ML-G -QKIGYSVFD   130

Consensus               ......K..K LN.N..L..S QL.L..LKN- ---------- -..IG..VF.   150


EST2 pep                FKQRLLKKFN NVL------- -PELYFMKFD VKSCYD                  157
Euplotes pep            YD-DVMKKYE EFVCKWKQVH CPKLFFATMD IEKCYD                  155
Trans of tetrahymen     NK-QISEKFA QFIEKWKNKG RPCLYYVTL- ------                  158

Consensus               .K-...KKF. .F..KWK..G .P.LYF.T.D ...CYD                  186
```

FIG. 53

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

FIG. 54

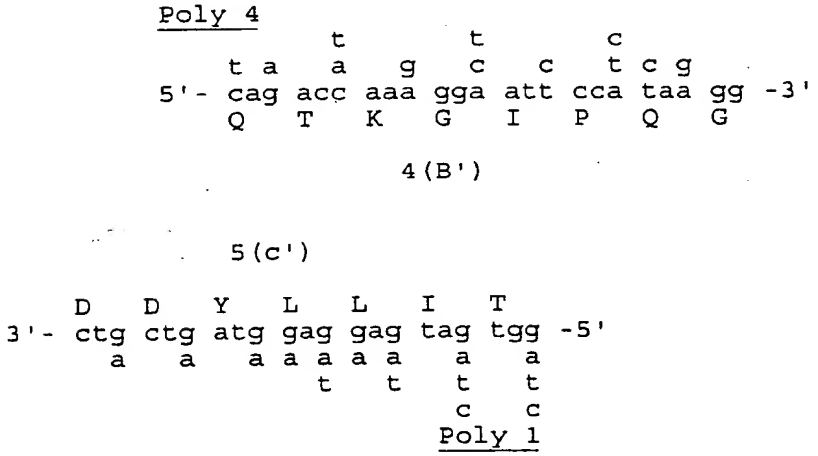


FIG. 56

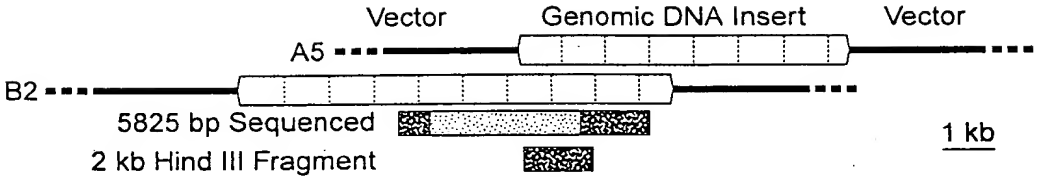


FIG. 55A

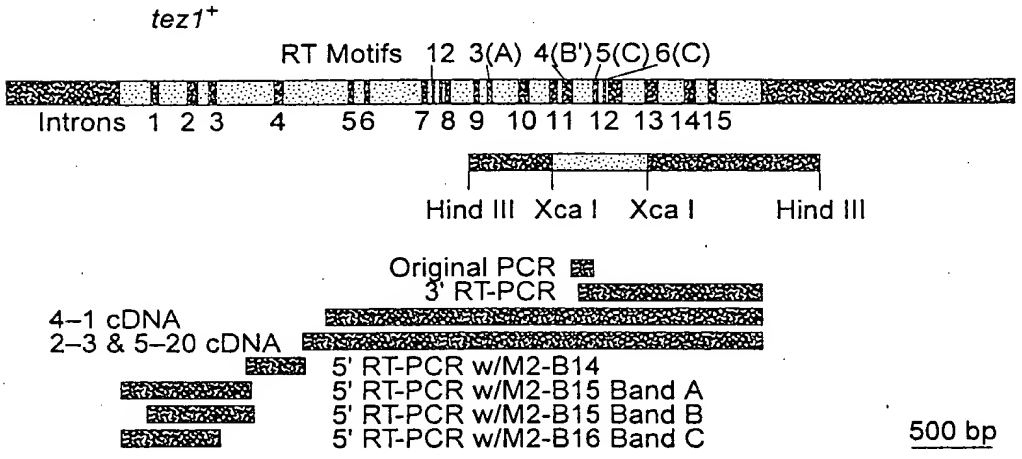


FIG. 55B

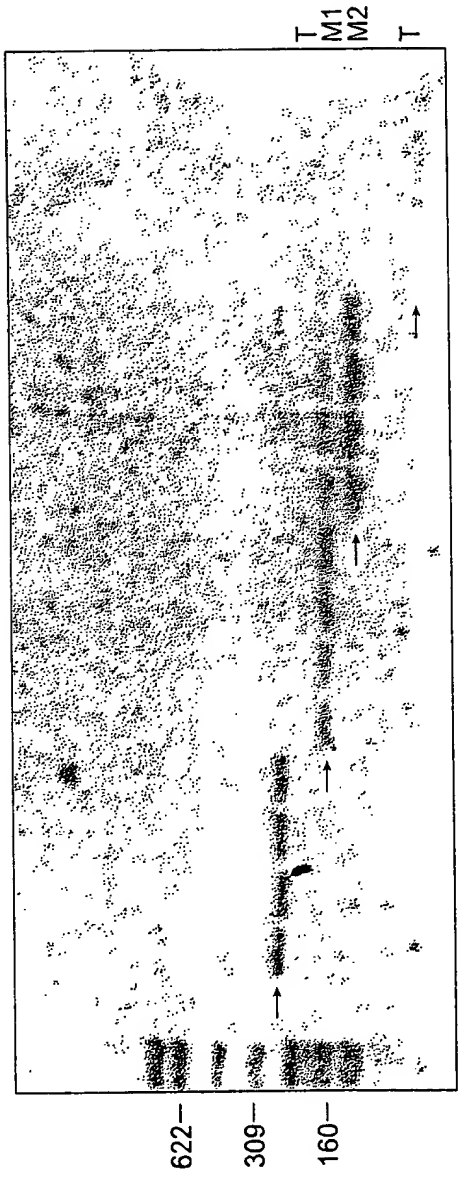


FIG. 57

```
Ot              LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Ea_p123     KGIPQGLCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2           SILSSFLCHFYMEDLIDEYLSFTKKK--------GSVLLRVV
Sc_p103     DGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS------QDTLILKLADDFLIIS
                    .*  .       * .    *              .....

Q   K   V   G   I   P   Q   G
caa aaa gtt ggt atc cct cag gg.....    <---Actual Genomic Sequence.

Poly 4
      t       t       c
t a   a   g   c   c   t c g
cag acc aaa gga att cca taa gg ---->

 ag acc aaa gga att cca tca ggC TCA ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
 tc tgg ttt cct taa ggt agt ccG AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC

       K   G   I   P   S   G   S   I   L   S   S   F   L   C   H   F   Y   M
```

FIG. 58A

```
GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT

E   D   L   I   D   E   Y   L   S   F   T   K   K   K   G   S   V   L   L   R


GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V   V   D   D   Y   L   L   I   T

  <---- ctg ctg atg gag gag tag tgg
          a   a   a a a a a   a   a
                      t   t   t   t
                              c   c
                            Poly 1

   .....gac gat ttc ctc ttt ata aca.....  <---Actual Genomic Sequence
        D   D   F   L   F   I   T
```

FIG. 58B

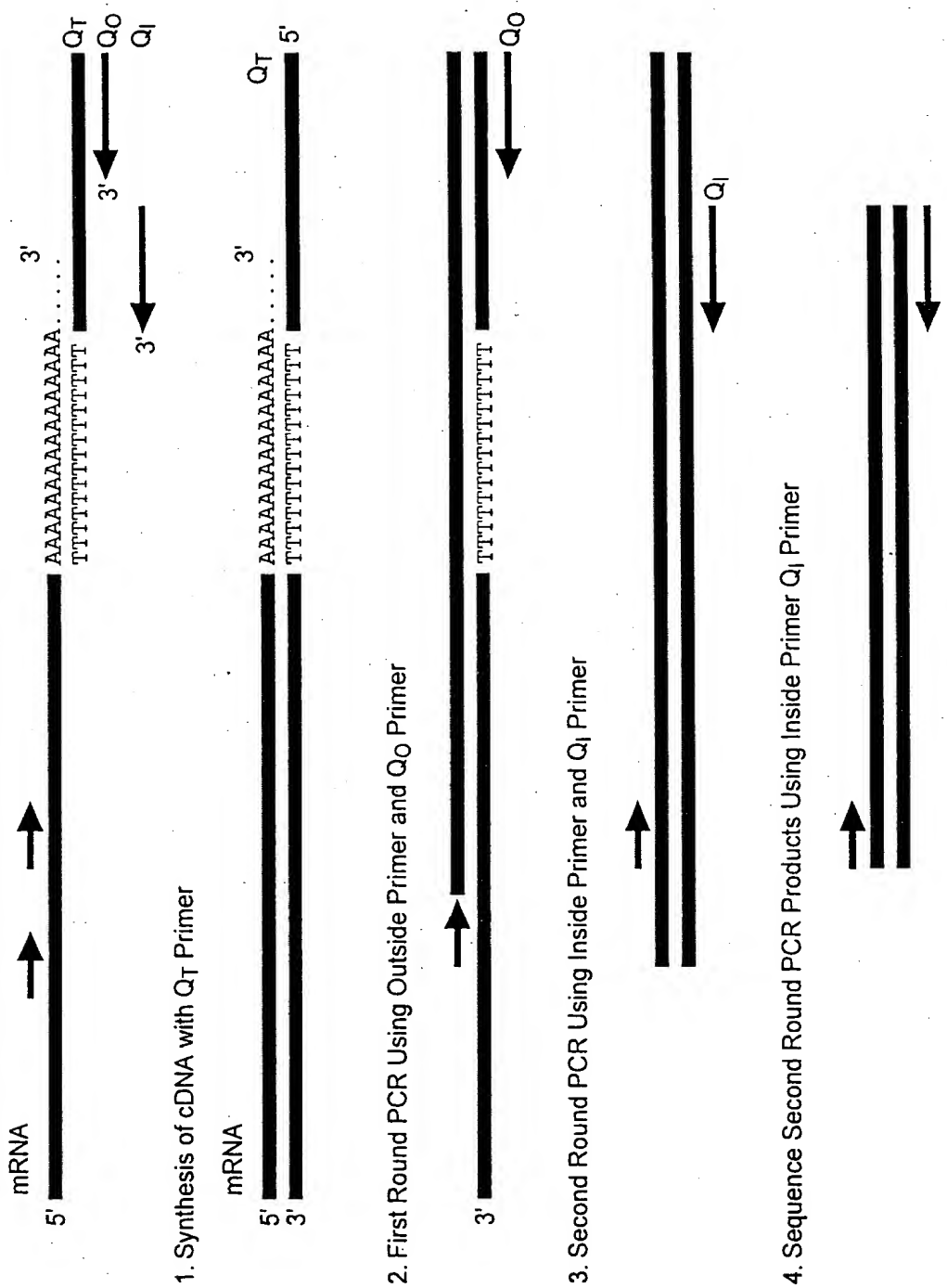


FIG. 59

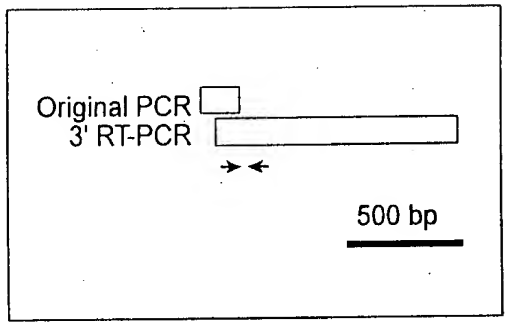
A. Genomic Libraries

Size Selected Libraries from P. Nurese
- 3~4 kb
- 5~6 kb
- 7~8 kb
- 11~12 kb

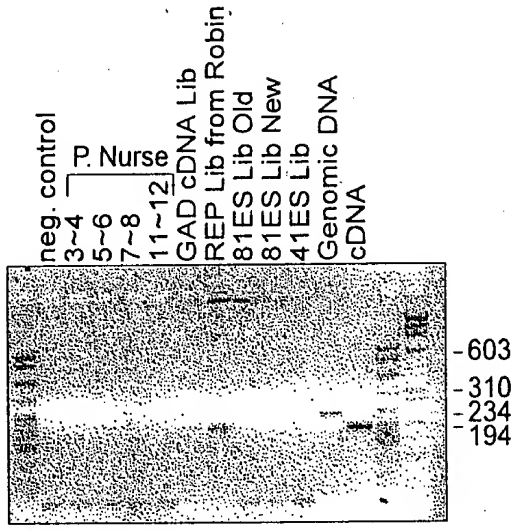
Libraries from J.A. Wise
- Sau 3a Partial Digest
- Hind III Partial Digest

cDNA Libraries
- GAD (Gal Activation Domain) Library
- REP Library from R. Allshire
- REP81ES Library (old)
- REP81ES Library (new)
- REP41ES Library

B.



C.



D.

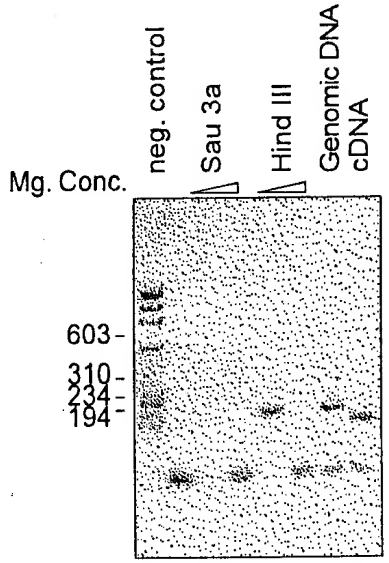


FIG. 60

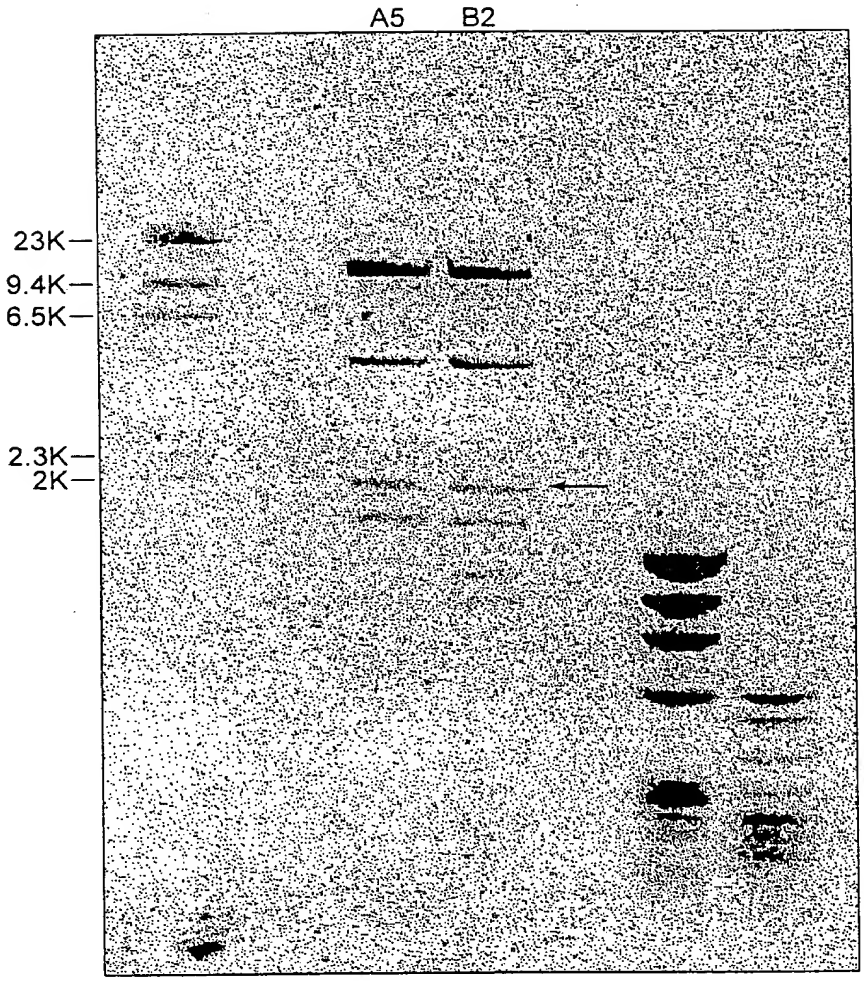


FIG. 61

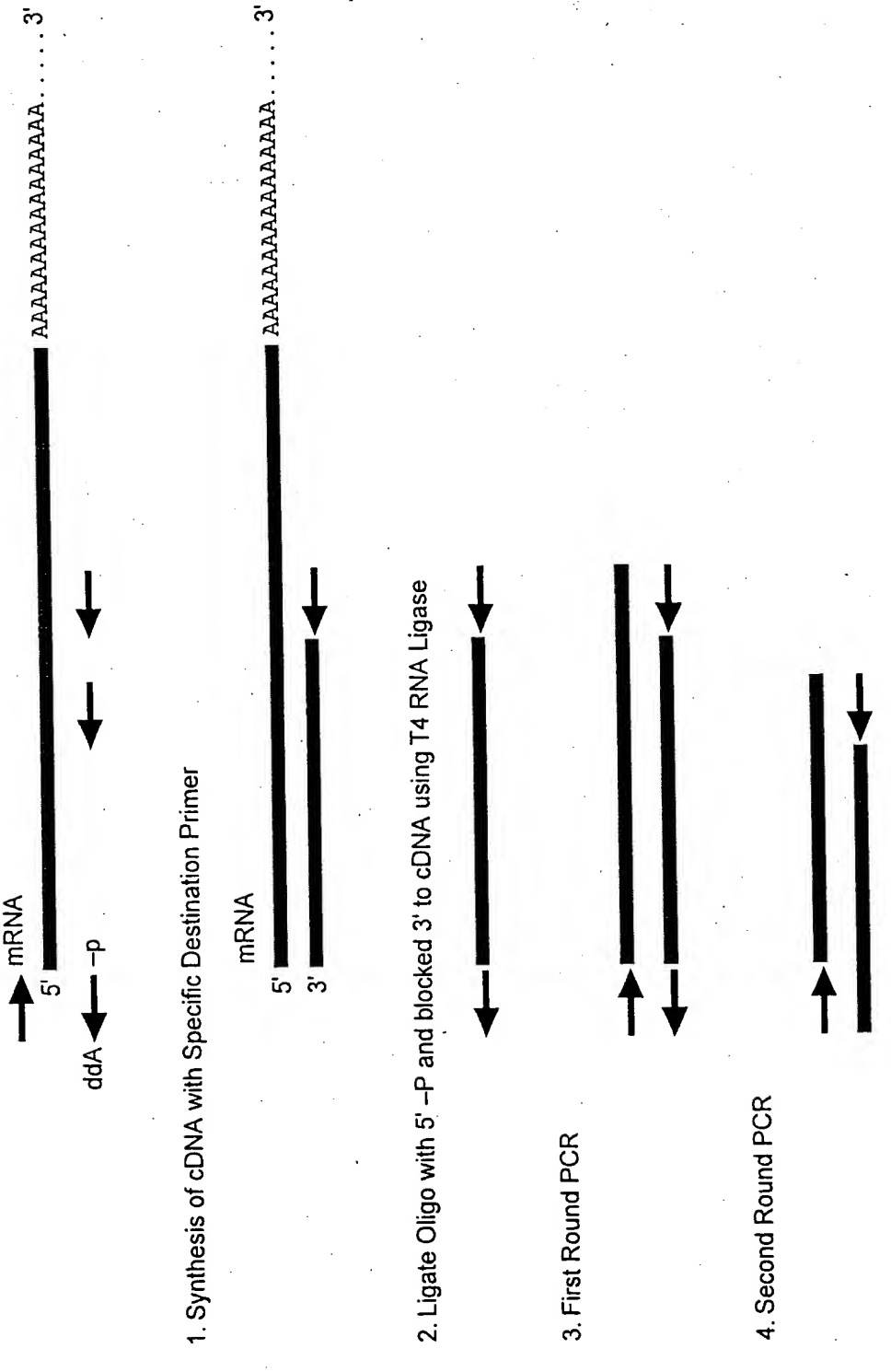


FIG. 62

```
                                   Motif O
S.p. Tez1p (429). WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIW ...(35)...
S.c. Est2p (366). WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTW ...(35)...
E.a. p123  (441). WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIW ...(35)...
                   *              *** **          * *   *

                    Motif 1    Motif 2     K
                   p hh h K     hR h       R
S.p. Tez1p  AVIRLLPKK--NTFRLITN-LRKRF ...(61)...
S.c. Est2p  SKMRIIPKKSNNEFRIIAIPCRGAD ...(62)...
E.a. p123   GKLRLIPKK--TTFRPIMTFNKKIV ...(61)...
              *  ***    ** *
                   Motif 3(A) AF
                      h  hDh  GY  h
S.p. Tez1p  KKYFVRIDIKSCYDRIKQDLMFRIVK ...(89)...
S.c. Est2p  ELYFMKFDVKSCYDSIPRMECMRILK ...(75)...
E.a. p123   KLFFATMDIEKCYDSVNREKLSTFLK ...(107)...
              *   *   ***          *
                    Motif 4(B')
                        hPQG   pP hh   h
S.p. Tez1p  YLQKVGIPQGSILSSFLCHFYMEDLIDEYLSF ...(6)...
S.c. Est2p  YIREDGLFQGSSLSAPIVDLVYDDLLEFYSEF ...(8)...
E.a. p123   YKQTKGIPQGLCVSSILSSFYYATLEESSLGF ...(14)...
            *    *  **   *          *     *
                   Y Motif 5(C)                    Motif 6(D)
                    h  F DDhhh                     Gh h  cK  h
S.p. Tez1p  VLLRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENS .(205)
S.c. Est2p  LILKLADDFLIISTDQQQVINIKKLAMGGFQKYNAKANRDKILAVSSQS .(173)
E.a. p123   LLMRLTDDYLLITTQENNAVLFIEKLINVSRENGFKFNMKKLQTSFPLS .(209)
                  ** * *                                  *       *
```

FIG. 63

A.

```
Sp_Tip1p    1   ---------MTEHHTPKSRILRFLENQYVYLCT  24
Sc_Est2p    1   ----------------------------MKILFEF   7
Ea_p123     1   MEVDVDNQADNHGIHSALKTCEEIKEAKTLYSW  33

Sp_Tip1p   25   LNDYVQLVLRGSPASSYSNICERLRSDVQTSFS  57
Sc_Est2p    8   IQDKLDIDLQTN--STYK---ENLKCGHFNGLD  35
Ea_p123    34   IQKVIRCRNQSQ--SHYK---DLEDIKIFAQTN  61

Sp_Tip1p   58   IFLHSTVVGFDSKPDEGVQFSSPKCSQSELIAN  90
Sc_Est2p   36   EILTTCFALPNSR-KIALPCLPGDLSHKAVIDH  67
Ea_p123    62   IVATPRDYNEEDFKVIARKEVFSTGLMIELIDK  94

Sp_Tip1p   91   VVKQMFDESFERRR-NLLMKGFSMNHEDFRAMH 122
Sc_Est2p   68   CIIYLLTGELYN---NVLTFGYKIARNED---- 93
Ea_p123    95   CLVELLSSSDVSDRQKLQCFGFQLKGNQ----- 122

Sp_Tip1p  123   VNGVQNDLVSTFPNYLISILESKNWQLLLEIIG 155
Sc_Est2p   94   ---VNNSLFCHSANVNVTLLKGAAWKMFHSLVG 123
Ea_p123   123   ---LAKTHLLTALSTQKQYFFQDEWNQVRAMIG 152

Sp_Tip1p  156   SDAMHYLLSKGSIFEALPNDNYLQISGIPLFKN 188
Sc_Est2p  124   TYAFVDLLINYTVIQFN-GQFFTQIVGNRCNEP 155
Ea_p123   153   NELFRHLYTKYLIFQRTSEGTLVQFCGNNVFDH 185

Sp_Tip1p  189   NVFEETVSKKRKRTIETSITQN---KSARKEVS 218
Sc_Est2p  156   HLPPKWVQ--RSSSSSATAAQI---KQLTEPVT 183
Ea_p123   186   LKVNDKFDK-KQKGGAADMNEPRCCSTCKYNVK 217
```

FIG. 64A

A.

```
Sp_Tip1p  219  WNSISISRFSIFYRSSYKKFKQDLYFNLHSICD  251
Sc_Est2p  184  N----------------KQFLHKLNINSSSFFP  200
Ea_p123   218  NEK--DHFLNNINVPNWNNMKSRTRIFYCTHFN  248

Sp_Tip1p  252  RNTVHMWLQWIFPRQFGLINAFQVKQLHKVIPL  284
Sc_Est2p  201  -------YSKILPSSS---SIKKLTDLREAIFP  223
Ea_p123   249  R------NNQFFKKHEFVSNKNNISAMDRAQTI  275

Sp_Tip1p  285  VS----QSTVVPKRLLKVYPLIEQTAKRLHRIS  313
Sc_Est2p  224  TN----LVKIPQRLKVRINLTLQKLLKRHKRLN  252
Ea_p123   276  FTNIFRFNRIRKKLKDKVIEKIAYMLEKVKDFN  308

Sp_Tip1p  314  LSKVYNHYCPYID-THDDEKILSYSLKPNQ---  342
Sc_Est2p  253  YVSILNSICPPLEGTVLDLSHLSRQSPKER---  282
Ea_p123   309  FNYYLTKSCPLPENWRERKQKIENLINKTREEK  341

Sp_Tip1p  343  ----------------VFAFLRSILVRVFPKLI  359
Sc_Est2p  283  ----------------VLKFIIVILQKLLPQEM  299
Ea_p123   342  SKYYEELFSYTTDNKCVTQFINEFFYNILPKDF  374

Sp_Tip1p  360  WGNQRIFEIILKDLETFLKLSRYESFSLHYLMS  392
Sc_Est2p  300  FGSKKNKGKIIKNLNLLLSLPLNGYLPFDSLLK  332
Ea_p123   375  LTG-RNRKNFQKKVKKYVELNKHELIHKNLLLE  406

Sp_Tip1p  393  NIKISEIEWLVLGKRSNAKMCLSDFEKRKQIFA  425
Sc_Est2p  333  KLRLKDFRWLFIS---DIWFTKHNFENLNQLAI  362
Ea_p123   407  KINTREISWMQVETS-AKHFYYFDHEN-IYVLW  437
```

FIG. 64B

A.

```
Sp_Tip1p  426  EFIYWLYNSFIIPILQSFFYITESSDLRNRTVY  458
Sc_Est2p  363  CFISWLFRQLIPKIIQTFFYCTEISSTVT-IVY  394
Ea_p123   438  KLLRWIFEDLVVSLIRCFFYVTEQQKSYSKTYY  470

Sp_Tip1p  459  FRKDIWKLLCRPFITSMKMEAFEKINENNVRMD  491
Sc_Est2p  395  FRHDTWNKLITPFIVEYFKTYLVENNVCRNHNS  427
Ea_p123   471  YRKNIWDVIMKMSIADLKKETLAEVQEKEVEEW  503

Sp_Tip1p  492  TQKTTLPPAVIRLLPKK--NTFRLITNLRKRFL  522
Sc_Est2p  428  YTLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD  460
Ea_p123   504  KKSLGFAPGKLRLIPKK--TTFRPIMTFNKKIV  534

Sp_Tip1p  523  IKMGSNKKMLVSTNQTLRPVASILKHLINE---  552
Sc_Est2p  461  EEE--FTIYKENHKNAIQPTQKILEYLRNKRPT  491
Ea_p123   535  NSD--RKTTKLTTNTKLLNSHLMLKTLKNR-MF  564

Sp_Tip1p  553  ESSGIPFNLEVYMKLLTFKKDLLKHRMFGR-KK  584
Sc_Est2p  492  SFTKIYSPTQIADRIKEFKQRLLKKFNNVLPEL  524
Ea_p123   565  KDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL  597

Sp_Tip1p  585  YFVRIDIKSCYDRIKQDLMFRIVKKKLKDPE-F  616
Sc_Est2p  525  YFMKFDVKSCYDSIPRMECMRILKDALKNENGF  557
Ea_p123   598  FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDF  630

Sp_Tip1p  617  VIRKYATIHATSDRATKN----------------  634
Sc_Est2p  558  FVRSQYFFNTNTG-------------------  570
Ea_p123   631  WIMTAQILKRKNNIVIDSKNFRKKEMKDYFRQK  663
```

FIG. 64C

A.

```
Sp_Tip1p  635  FVSEAFSYFDMVPFEKVVQLLS--MKTSDTLFV  665
Sc_Est2p  571  ----------VLKLFNVVNASR--VPKPYELYI  591
Ea_p123   664  FQKIALEGGQYPTLFSVLENEQNDLNAKKTLIV  696

Sp_Tip1p  666  DFVDYWTKSSSEIFKMLKEHLSGHIVKIGNSQY  698
Sc_Est2p  592  DNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY  624
Ea_p123   697  EAKQRNYFKKDNLLQPVINICQYNYINFNGKFY  729

Sp_Tip1p  699  LQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTK  731
Sc_Est2p  625  IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKA  657
Ea_p123   730  KQTKGIPQGLCVSSILSSFYYATLEESSLGFLR  762

Sp_Tip1p  732  KKG--------SVLLRVVDDFLFITVNKKDAKK  756
Sc_Est2p  658  SPSQD------TLILKLADDFLIISTDQQQVIN  684
Ea_p123   763  DESMNPENPNVNLLMRLTDDYLLITTQENNAVL  795

Sp_Tip1p  757  FLNLSLRGFEKHNFSTSLEKTVINFENSNG---  786
Sc_Est2p  685  IKKLAMGGFQKYNAKANRDKILAVSSQSD----  713
Ea_p123   796  FIEKLINVSRENGFKFNMKKLQTSFPLSPSKFA  828

Sp_Tip1p  787  ---IINNTFFNESKKRMPFFGFSVNMRSLDTLL  816
Sc_Est2p  714  ----DDTVIQFCA---MHIFVKELEVWKHSSTM  739
Ea_p123   829  KYGMDSVEEQNIVQDYCDWIGISIDMKTLALMP  861

Sp_Tip1p  817  ACPKIDEALFNSTSVELTKHMGKSFFYKILRSS  849
Sc_Est2p  740  NNFHIRSKSSKGIFRSLIALFNTRISYKTIDTN  772
Ea_p123   862  NINLRIEGILCTLNLNMQTKKASMWLKKKLKSF  894
```

FIG. 64D

A.

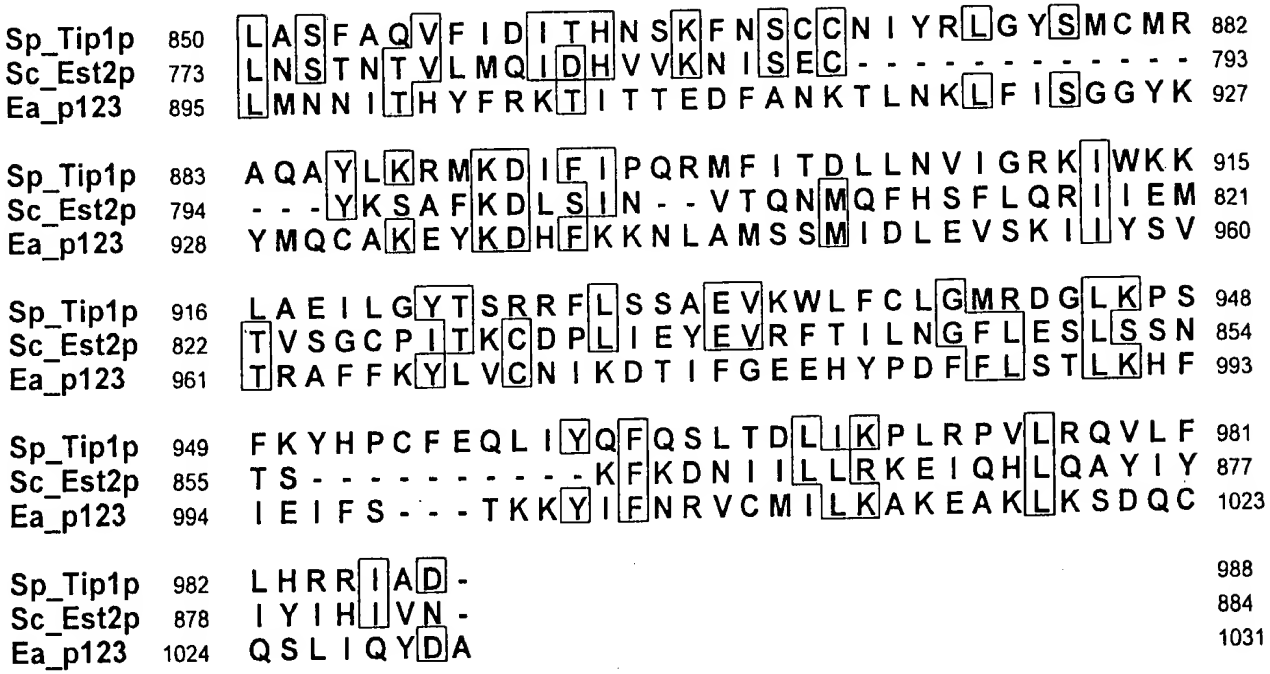


FIG. 64E

B.

```
Sp_Tip1p    1  ----------MTEHHTPKSRILRFLENQYVYLCT 24
Sc_Est2p    1  ----------------------------MKILFEF  7
Ea_p123     1  MEVDVDNQADNHGIHSALKTCEEIKEAKTLYSW 33

Sp_Tip1p   25  LNDYVQLVLRGSPASSYSNICERLRSDVQTSFS 57
Sc_Est2p    8  IQDKLDIDLQTN--STYK---ENLKCGHFNGLD 35
Ea_p123    34  IQKVIRCRNQSQ--SHYK---DLEDIKIFAQTN 61

Sp_Tip1p   58  IFLHSTVVGFDSKPDEGVQFSSPKCSQSELIAN 90
Sc_Est2p   36  EILTTCFALPNSR-KIALPCLPGDLSHKAVIDH 67
Ea_p123    62  IVATPRDYNEEDFKVIARKEVFSTGLMIELIDK 94

Sp_Tip1p   91  VVKQMFDESFERRR-NLLMKGFSMNHEDFRAMH 122
Sc_Est2p   68  CIIYLLTGELYN---NVLTFGYKIARNED---- 93
Ea_p123    95  CLVELLSSSDVSDRQKLQCFGFQLKGNQ----- 122

Sp_Tip1p  123  VNGVQNDLVSTFPNYLISILESKNWQLLLEIIG 155
Sc_Est2p   94  ---VNNSLFCHSANVNVTLLKGAAWKMFHSLVG 123
Ea_p123   123  ---LAKTHLLTALSTQKQYFFQDEWNQVRAMIG 152

Sp_Tip1p  156  SDAMHYLLSKGSIFEALPNDNYLQISGIPLFKN 188
Sc_Est2p  124  TYAFVDLLINYTVIQFN-GQFFTQIVGNRCNEP 155
Ea_p123   153  NELFRHLYTKYLIFQRTSEGTLVQFCGNNVFDH 185

Sp_Tip1p  189  NVFEETVSKKRKRTIETSITQN---KSARKEVS 218
Sc_Est2p  156  HLPPKWVQ--RSSSSSATAAQI---KQLTEPVT 183
Ea_p123   186  LKVNDKFDK-KQKGGAADMNEPRCCSTCKYNVK 217
```

FIG. 64F

B.

```
Sp_Tip1p  219  WNSISISRFSIFYRSSYKKFKQDLYFNLHSICD  251
Sc_Est2p  184  N-----------------KQFLHKLNINSSSFFP  200
Ea_p123   218  NEK--DHFLNNINVPNWNNMKSRTRIFYCTHFN  248

Sp_Tip1p  252  RNTVHMWLQWIFPRQFGLINAFQVKQLHKVIPL  284
Sc_Est2p  201  -------YSKILPSSS---SIKKLTDLREAIFP  223
Ea_p123   249  R------NNQFFKKHEFVSNKNNISAMDRAQTI  275

Sp_Tip1p  285  VS----QSTVVPKRLLKVYPLIEQTAKRLHRIS  313
Sc_Est2p  224  TN----LVKIPQRLKVRINLTLQKLLKRHKRLN  252
Ea_p123   276  FTNIFRFNRIRKKLKDKVIEKIAYMLEKVKDFN  308

Sp_Tip1p  314  LSKVYNHYCPYID-THDDEKILSYSLKPNQ---  342
Sc_Est2p  253  YVSILNSICPPLEGTVLDLSHLSRQSPKER---  282
Ea_p123   309  FNYYLTKSCPLPENWRERKQKIENLINKTREEK  341

Sp_Tip1p  343  ----------------VFAFLRSILVRVFPKLI  359
Sc_Est2p  283  ----------------VLKFIIVILQKLLPQEM  299
Ea_p123   342  SKYYEELFSYTTDNKCVTQFINEFFYNILPKDF  374

Sp_Tip1p  360  WGNQRIFEIILKDLETFLKLSRYESFSLHYLMS  392
Sc_Est2p  300  FGSKKNKGKIIKNLNLLLSLPLNGYLPFDSLLK  332
Ea_p123   375  LTG-RNRKNFQKKVKKYVELNKHELIHKNLLLE  406

Sp_Tip1p  393  NIKISEIEWLVLGKRSNAKMCLSDFEKRKQIFA  425
Sc_Est2p  333  KLRLKDFRWLFIS---DIWFTKHNFENLNQLAI  362
Ea_p123   407  KINTREISWMQVETS-AKHFYYFDHEN-IYVLW  437
```

FIG. 64G

B.

```
Sp_Tip1p  426  EFIYWLYNSFIIPILQSFFYITESSDLRNRTVY  458
Sc_Est2p  363  CFISWLFRQLIPKIIQTFFYCTEISSTVT-IVY  394
Ea_p123   438  KLLRWIFEDLVVSLIRCFFYVTEQQKSYSKTYY  470

Sp_Tip1p  459  FRKDIWKLLCRPFITSMKMEAFEKINENNVRMD  491
Sc_Est2p  395  FRHDTWNKLITPFIVEYFKTYLVENNVCRNHNS  427
Ea_p123   471  YRKNIWDVIMKMSIADLKKETLAEVQEKEVEEW  503

Sp_Tip1p  492  TQKTTLPPAVIRLLPKK--NTFRLITNLRKRFL  522
Sc_Est2p  428  YTLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD  460
Ea_p123   504  KKSLGFAPGKLRLIPKK--TTFRPIMTFNKKIV  534

Sp_Tip1p  523  IKMGSNKKMLVSTNQTLRPVASILKHLINE---  552
Sc_Est2p  461  EEE--FTIYKENHKNAIQPTQKILEYLRNKRPT  491
Ea_p123   535  NSD--RKTTKLTTNTKLLNSHLMLKTLKNR-MF  564

Sp_Tip1p  553  ESSGIPFNLEVYMKLLTFKKDLLKHRMFGR-KK  584
Sc_Est2p  492  SFTKIYSPTQIADRIKEFKQRLLKKFNNVLPEL  524
Ea_p123   565  KDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL  597

Sp_Tip1p  585  YFVRIDIKSCYDRIKQDLMFRIVKKKLKDPE-F  616
Sc_Est2p  525  YFMKFDVKSCYDSIPRMECMRILKDALKNENGF  557
Ea_p123   598  FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDF  630

Sp_Tip1p  617  VIRKYATIHATSDRATKN----------------  634
Sc_Est2p  558  FVRSQYFFNTNTG-------------------  570
Ea_p123   631  WIMTAQILKRKNNIVIDSKNFRKKEMKDYFRQK  663
```

FIG. 64H

B.

```
Sp_Tip1p  635  FVSEAFSYFDMVPFEKVVQLLS--MKTSDTLFV  665
Sc_Est2p  571  ----------VLKLFNVVNASR--VPKPYELYI  591
Ea_p123   664  FQKIALEGGQYPTLFSVLENEQNDLNAKKTLIV  696

Sp_Tip1p  666  DFVDYWTKSSSEIFKMLKEHLSGHIVKIGNSQY  698
Sc_Est2p  592  DNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY  624
Ea_p123   697  EAKQRNYFKKDNLLQPVINICQYNYINFNGKFY  729

Sp_Tip1p  699  LQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTK  731
Sc_Est2p  625  IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKA  657
Ea_p123   730  KQTKGIPQGLCVSSILSSFYYATLEESSLGFLR  762

Sp_Tip1p  732  KKG--------SVLLRVVDDFLFITVNKKDAKK  756
Sc_Est2p  658  SPSQD------TLILKLADDFLIISTDQQQVIN  684
Ea_p123   763  DESMNPENPNVNLLMRLTDDYLLITTQENNAVL  795

Sp_Tip1p  757  FLNLSLRGFEKHNFSTSLEKTVINFENSNG---  786
Sc_Est2p  685  IKKLAMGGFQKYNAKANRDKILAVSSQSD----  713
Ea_p123   796  FIEKLINVSRENGFKFNMKKLQTSFPLSPSKFA  828

Sp_Tip1p  787  ---IINNTFFNESKKRMPFFGFSVNMRSLDTLL  816
Sc_Est2p  714  ----DDTVIQFCA---MHIFVKELEVWKHSSTM  739
Ea_p123   829  KYGMDSVEEQNIVQDYCDWIGISIDMKTLALMP  861

Sp_Tip1p  817  ACPKIDEALFNSTSVELTKHMGKSFFYKILRSS  849
Sc_Est2p  740  NNFHIRSKSSKGIFRSLIALFNTRISYKTIDTN  772
Ea_p123   862  NINLRIEGILCTLNLNMQTKKASMWLKKKLKSF  894
```

FIG. 64I

B.

```
Sp_Tip1p   850  LASFAQVFIDITHNSKFNSCCNIYRLGYSMCMR  882
Sc_Est2p   773  LNSTNTVLMQIDHVVKNISEC-------------  793
Ea_p123    895  LMNNITHYFRKTITTEDFANKTLNKLFISGGYK  927

Sp_Tip1p   883  AQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKK  915
Sc_Est2p   794  ---YKSAFKDLSIN--VTQNMQFHSFLQRIIEM  821
Ea_p123    928  YMQCAKEYKDHFKKNLAMSSMIDLEVSKIIYSV  960

Sp_Tip1p   916  LAEILGYTSRRFLSSAEVKWLFCLGMRDGLKPS  948
Sc_Est2p   822  TVSGCPITKCDPLIEYEVRFTILNGFLESLSSN  854
Ea_p123    961  TRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHF  993

Sp_Tip1p   949  FKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLF  981
Sc_Est2p   855  TS----------KFKDNIILLRKEIQHLQAYIY  877
Ea_p123    994  IEIFS---TKKYIFNRVCMILKAKEAKLKSDQC 1023

Sp_Tip1p   982  LHRRIAD-                           988
Sc_Est2p   878  IYIHIVN-                           884
Ea_p123   1024  QSLIQYDA                          1031
```

FIG. 64J

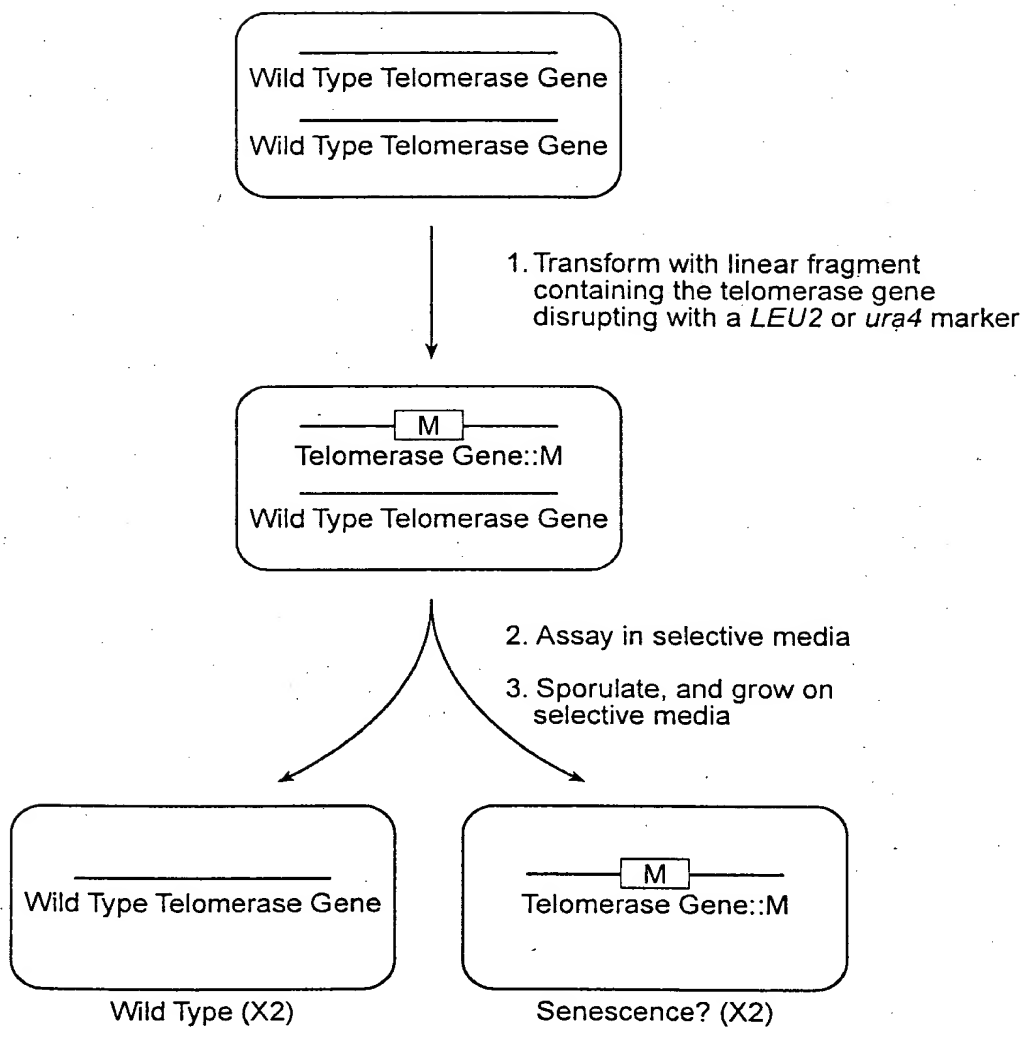


FIG. 65

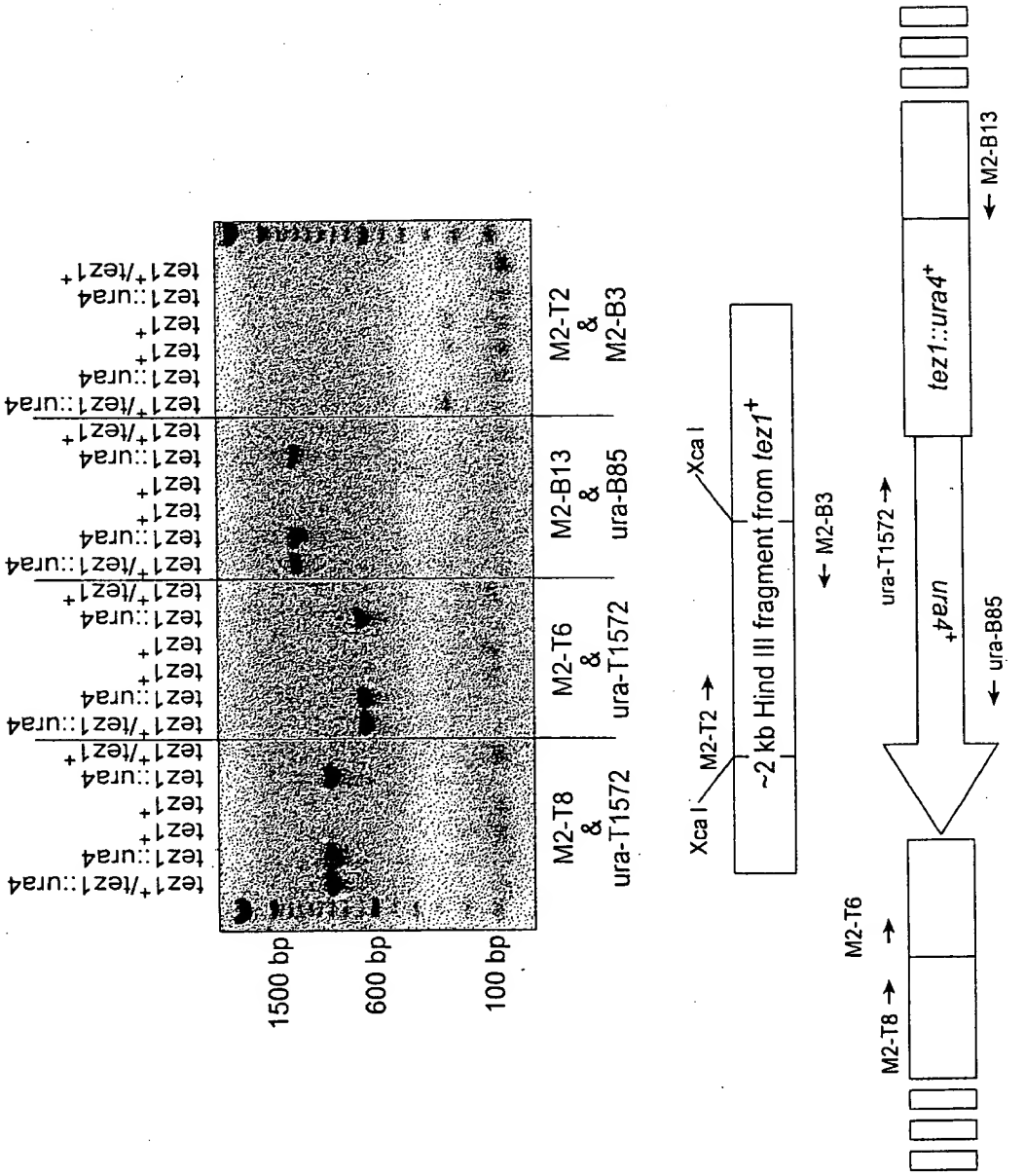


FIG. 66

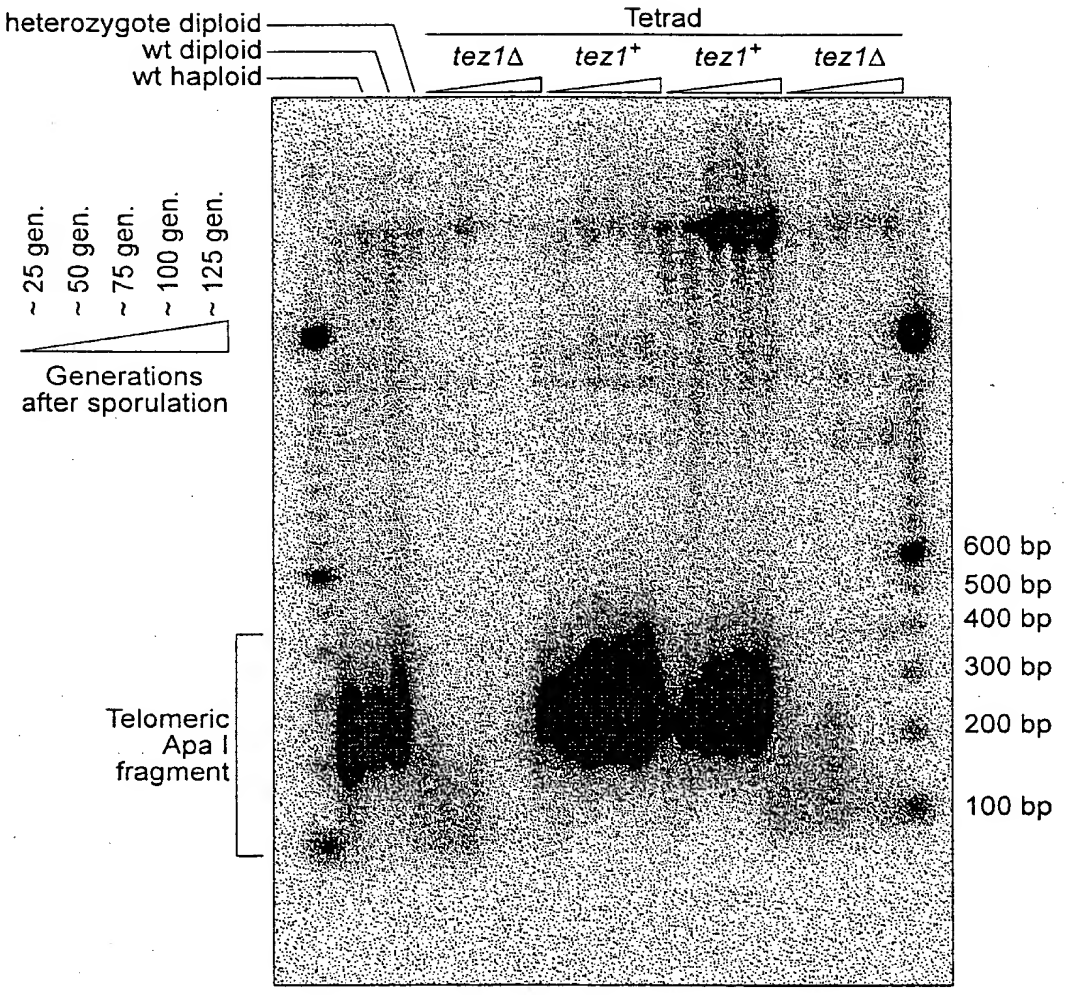


FIG. 67

```
                                1
                           met ser val tyr val val glu leu leu
GCCAAGTTCCTGCACTGGCTG      ATG AGT GTG TAC GTC GTC GAG CTG CTC

 10                                                  20
arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

                        30
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

 40                                                  50
gly ile arg gln his leu lys arg val gln leu arg glu leu ser
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

                        60
glu ala glu val arg gln his arg glu ala arg pro ala leu leu
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

 70                                                  80
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

                        90
ile val asn met asp tyr val val gly ala arg thr phe arg arg
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

100                                                 110
glu lys     ala glu arg leu thr ser arg val lys ala leu phe
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

                       120
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly
AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG GGC

130                                                 140
ala ser val leu gly leu asp asp ile his arg ala trp arg thr
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

                       150
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG TAC

160                                                 170
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

                       180
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC

190                                                 200
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG
```

FIG. 68A

```
                    210
gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220                                         230
gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

                    240
leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250                                         260
arg asp gly leu leu leu arg leu val asp asp phe leu leu val
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

                    270
thr pro his leu thr his ala lys thr phe leu arg thr leu val
ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280                                         290
arg gly val pro glu tyr gly cys val val asn leu arg lys thr
CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

                    300
val val asn phe pro val glu asp glu ala leu gly gly thr ala
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310                                         320
phe val gln met pro ala his gly leu phe pro trp cys gly leu
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

                    330
leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340                                         350
tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

                    360
phe lys ala gly arg asn met arg arg lys leu phe gly val leu
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370                                         380
arg leu lys cys his ser leu phe leu asp leu gln val asn ser
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

                    390
leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400                                         410
ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG
```

FIG. 68B

```
                    420
gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430                                             440
arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

                    450
val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460                                             470
arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

                    480
thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490                                             500
pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

                    510
pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520                                             530
his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

                    540
thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550                                             560
arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564
OP
TGA GTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC

CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTTCACTTCCCCAC

AGGCTGGCGTTCGGTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACT

CCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCTTCGCCCTGCCTTCC

TTTGCCTTCCACCCCCACCATTCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTTTGGG

AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGG

GTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTT

TTTCAGTTTTGGAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 68C

```
Motif -1
Ep p123          ...LVVSLIRCFFYVTEQQKSYSKT...
Sp Tez1          ...FIIPILQSFFYITESSDLRNRT...
Sc Est2          ...LIPKIIQTFFYCTEISSTVTIV...
Hs TCP1          ...YVVELLRSFFYVTETTFQKNRL...
consensus                     FFY TE

                                                  K
Motif 0                    p hhh  K     hR h      R
Ep p123          ...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...
Sp Tez1          ...QKTTLPPAVIRLLPKKN--TFRLITNLRKRFL...
Sc Est2          ...TLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD...
Hs TCP1          ...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...
consensus                      R  PK      R I

                                 AF
Motif A                    h  hDh  GY  h
Ep p123          ...PKLFFATMDIEKCYDSVNREKLSTFLK...
Sp Tez1          ...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
Sc Est2          ...PELYFMKFDVKSCYDSIPRMECMRILK...
Hs TCP1          ...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
consensus              F   D    YD

Motif B                         hPQG   pS hh
Ep p123          ...NGKFYKQTKGIPQGLCVSSILSSFYYA...
Sp Tez1          ...GNSQYLQKVGIPQGSILSSFLCHFYME...
Sc Est2          ...EDKCYIREDGLFQGSSLSAPIVDLVYD...
Hs TCP1          ...RATSYVQCQGIPQGSILSTLLCSLCYG...
consensus                     G  QG   S

                                Y
Motif C                    h  F DD hhh
Ep p123          ...PNVNLLMRLTDDYLLITTQENN...
Sp Tez1          ...KKGSVLLRVVDDFLFITVNKKD...
Sc Est2          ...SQDTLILKLADDFLIISTDQQQ...
Hs TCP1          ...RRDGLLLRLVDDFLLVTPHLTH...
consensus                      DD L

Motif D                     Gh h  cK
Ep p123          ...NVSRENGFKFNMKKL...
Sp Tez1          ...LNLSLRGFEKHNFST...
Sc Est2          ...KKLAMGGFQKYNAKA...
Hs TCP1          ...LRTLVRGVPEYGCVV...
consensus                   G
```

FIG. 69

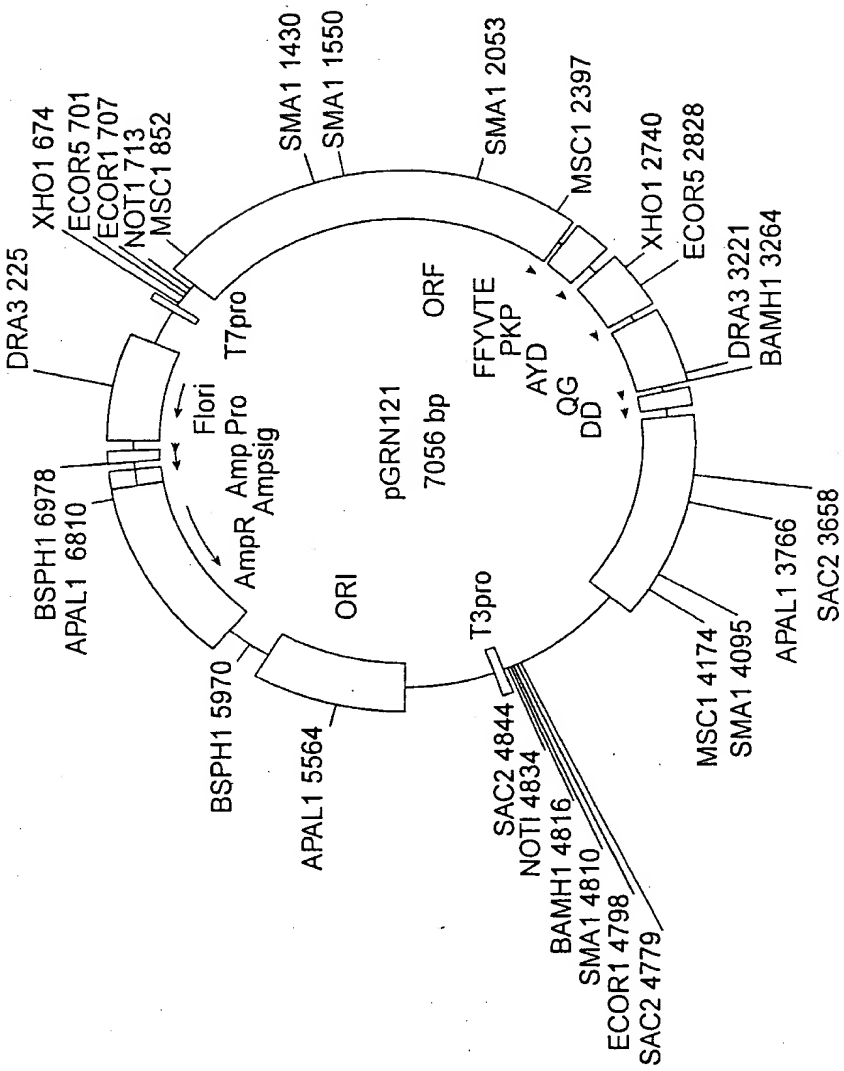


FIG. 70

```
   1  GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC
  51  CCGCGATGCC GCGCGCTCCC CGCTGCCGAG CCGTGCGCTC CCTGCTGCGC
 101  AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTTCGTGC GGCGCCTGGG
 151  GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG GGACCCGGCG GCTTTCCGCG
 201  CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC
 251  CCCGCCGCCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC
 301  CCGAGTGCTG CANANGCTGT GCGANCGCGG CGCGAANAAC GTGCTGGCCT
 351  TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCCCC CGAGGCCTTC
 401  ACCACCAGCG TGCGCAGCTA CCTGCCCAAC ACGGTGACCG ACGCACTGCG
 451  GGGGAGCGGG GCGTGGGGGC TGCTGCTGCG CCGCGTGGGC GACGACGTGC
 501  TGGTTCACCT GCTGGCACGC TGCGCGNTNT TTGTGCTGGT GGNTCCCAGC
 551  TGCGCCTACC ANGTGTGCGG GCCGCCGCTG TACCAGCTCG GCGCTGCNAC
 601  TCAGGCCCGG CCCCCGCCAC ACGCTANTGG ACCCGAANGC GTCTGGGATC
 651  CAACGGGCCT GGAACCATAG CGTCAGGGAG GCCGGGGTCC CCCTGGGCTG
 701  CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC
 751  GTTGCCCAAG AGGCCCAGGC GTGGCGCTGC CCCTGAGCCG GAGCGGACGC
 801  CCGTTGGGCA GGGGTCCTGG GCCCACCCGG GCAGGACGCC TGGACCGAGT
 851  GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCCG AAGAAGCCAC
 901  CTCTTTGGAG GGTGCGCTCT CTGGCACGCG CCACTCCCAC CCATCCGTGG
 951  GCCGCCAGCA CCACGCGGGC CCCCCATCCA CATCGCGGCC ACCACGTCCT
1001  GGGACACGCC TTGTCCCCCG GTGTACGCCG AGACCAAGCA CTTCCTCTAC
1051  TCCTCAGGCG ACAAGNACAC TGCGNCCCTC CTTCCTACTC AATATATCTG
1101  AGGCCCAGCC TGACTGGCGT TCGGGAGGTT CGTGGAGACA NTCTTTCTGG
1151  TTCCAGGCCT TGGATGCCAG GATTCCCCGC AGGTTGCCCC GCCTGCCCCA
1201  GCGNTACTGG CAAATGCGGC CCCTGTTTCT GGAGCTGCTT GGGAACCACG
1251  CGCAGTGCCC CTACGGGGTG TTCCTCAAGA CGCACTGCCC GCTGCGAGCT
1301  GCGGTCACCC CAGCAGCCGG TGTCTGTGCC CGGGAGAAGC CCCAGGGCTC
1351  TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCCGTCG CCTGGTGCAG
1401  CTGCTCCGCC AGCACAGCAG CCCCTGGCAG GTGTACGGCT TCGTGCGGGC
1451  CTGCCTGCGC CGGCTGGTGC CCCCAGGCCT CTGGGGCTCC AGGCACAACG
1501  AACGCCGCTT CCTCAGGAAC ACCAAGAAGT TCATCTCCCT GGGGAAGCAT
1551  GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TGCGGGACTG
1601  CGCTTGGCTG CGCAGGAGCC CAGGGGTTGG CTGTGTTCCG GCCGCAGAGC
1651  ACCGTCTGCG TGAGGAGATC CTGGCCAAGT TCCTGCACTG GCTGATGAGT
1701  GTGTACGTCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCAC
1751  GTTTCAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTC TGGAGCAAGT
1801  TGCAAAGCAT TGGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGCGGGAG
1851  CTGTCGGAAG CAGAGGTCAG GCAGCATCGG GAAGCCAGGC CCGCCCTGCT
1901  GACGTCCAGA CTCCGCTTCA TCCCCAAGCC TGACGGGCTG CGGCCGATTG
1951  TGAACATGGA CTACGTCGTG GGAGCCAGAA CGTTCCGCAG AGAAAAGAGG
2001  GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTCAGCG TGCTCAACTA
2051  CGAGCGGGCG CGGCGCCCCG GCCTCCTGGG CGCCTCTGTG CTGGGCCTGG
2101  ACGATATCCA CAGGGCCTGG CGCACCTTCG TGCTGCGTGT GCGGGCCCAG
2151  GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA
2201  CGACACCATC CCCCAGGACA GGCTCACGGA GGTCATCGCC AGCATCATCA
2251  AACCCCAGAA CACGTACTGC GTGCGTCGGT ATGCCGTGGT CCAGAAGGCC
2301  GCCCATGGGC ACGTCCGCAA GGCCTTCAAG AGCCACGTCT CTACCTTGAC
2351  AGACCTCCAG CCGTACATGC GACAGTTCGT GGCTCACCTG CAGGANAACA
2401  GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG
2451  GCCAGCAGTG GCCTCTTCGA CGTCTTCCTA CGCTTCATGT GCCACCACGC
```

FIG. 71A

```
2501  CGTGCGCATC AGGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG
2551  GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
2601  AACAAGCTGT TTGCGGGGAT TCGGCGGGAC GGGCTGCTCC TGCGTTTGGT
2651  GGATGATTTC TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
2701  TCAGGACCCT GGTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACTTG
2751  CGGAAGACAG TGGTGAACTT CCCTGTAGAA GACGAGGCCC TGGGTGGCAC
2801  GGCTTTTGTT CAGATGCCGG CCCACGGCCT ATTCCCCTGG TGCGGCCTGC
2851  TGCTGGATAC CCGGACCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
2901  CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG
2951  GAGGAACATG CGTCGCAAAC TCTTTGGGGT CTTGCGGCTG AAGTGTCACA
3001  GCCTGTTTCT GGATTTGCAG GTGAACAGCC TCCAGACGGT GTGCACCAAC
3051  ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTCACG CATGTGTGCT
3101  GCAGCTCCCA TTTCATCAGC AAGTTTGGAA GAACCCCACA TTTTTCCTGC
3151  GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
3201  AACGCAGGGA TGTCGCTGGG GGCCAAGGGC GCCGCCGGCC CTCTGCCCTC
3251  CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCCTGCTC AAGCTGACTC
3301  GACACCGTGT CACCTACGTG CCACTCCTGG GGTCACTCAG GACAGCCCAG
3351  ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC
3401  CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
3451  GGCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC
3501  GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCCACAC CCAGGCCCGC
3551  ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
3601  CCGGCTGAAG GCTGAGTGTC CGGCTGAGGC CTGAGCGAGT GTCCAGCCAA
3651  GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG
3701  CTCGGCTCCA CCCCAGGGCC AGCTTTTCCT CACCAGGAGC CCGGCTTCCA
3751  CTCCCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC
3801  GCCCTGCCCT CCTTTGCCTT CCACCCCCAC CATCCAGGTG GAGACCCTGA
3851  GAAGGACCCT GGGAGCTCTG GGAATTTGGA GTGACCAAAG GTGTGCCCTG
3901  TACACAGGCG AGGACCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
3951  GGGGGGAGGT GCTGTGGGAG TAAAATACTG AATATATGAG TTTTTCAGTT
4001  TTGAAAAAAA AAAAAAAAAA AAAAAAAAA
```

FIG. 71B

```
                GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCGATGCC
              1 ---------+---------+---------+---------+---------+---------+ 60
                CGTCGCGACGCAGGACGACGCGTGCACCCTTCGGGACCGGGGCCGGTGGGGGCGCTACGG

a                 A  A  L  R  P  A  A  H  V  G  S  P  G  P  G  H  P  R  D  A  -
b                  Q  R  C  V  L  L  R  T  W  E  A  L  A  P  A  T  P  A  M  P  -
c                   S  A  A  S  C  C  A  R  G  K  P  W  P  R  P  P  P  R  C  R  -

                GCGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCT
             61 ---------+---------+---------+---------+---------+---------+ 120
                CGCGCGAGGGGCGACGGCTCGGCACGCGAGGGACGACGCGTCGGTGATGGCGCTCCACGA

a                 A  R  S  P  L  P  S  R  A  L  P  A  A  Q  P  L  P  R  G  A  -
b                  R  A  P  R  C  R  A  V  R  S  L  L  R  S  H  Y  R  E  V  L  -
c                   A  L  P  A  A  E  P  C  A  P  C  C  A  A  T  T  A  R  C  C  -

                GCCGCTGGCCACGTTCGTGCGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGG
            121 ---------+---------+---------+---------+---------+---------+ 180
                CGGCGACCGGTGCAAGCACGCCGCGGACCCCGGGGTCCCGACCGCCGACCACGTCGCGCC

a                 A  A  G  H  V  R  A  A  P  G  A  P  G  L  A  A  G  A  A  R  -
b                  P  L  A  T  F  V  R  R  L  G  P  Q  G  W  R  L  V  Q  R  G  -
c                   R  W  P  R  S  C  G  A  W  G  P  R  A  G  G  W  C  S  A  G  -

                GGACCCGGCGGCTTTCCGCGCGNTGGTGGCCCANTGCNTGGTGTGCGTGCCCTGGGANGN
            181 ---------+---------+---------+---------+---------+---------+ 240
                CCTGGGCCGCCGAAAGGCGCGCNACCACCGGGTNACGNACCACACGCACGGGACCCTNCN

a                 G  P  G  G  F  P  R  ?  G  G  P  ?  ?  G  V  R  A  L  G  ?  -
b                  D  P  A  A  F  R  A  ?  V  A  ?  C  ?  V  C  V  P  W  ?  ?  -
c                   T  R  R  L  S  A  R  W  W  P  ?  A  W  C  A  C  P  G  ?  ?  -

                ANGGCNGCCCCCCGCCGCCCCCTCCTTCCGCCAGGTGTCCTGCCTGAANGANCTGGTGGC
            241 ---------+---------+---------+---------+---------+---------+ 300
                TNCCGNCGGGGGGCGGCGGGGGAGGAAGGCGGTCCACAGGACGGACTTNCTNGACCACCG

a                 ?  A  A  P  R  R  P  L  L  P  P  G  V  L  P  E  ?  ?  G  G  -
b                  ?  ?  P  P  A  A  P  S  F  R  Q  V  S  C  L  ?  ?  L  V  A  -
c                   G  ?  P  P  P  P  P  P  S  A  R  C  P  A  *  ?  ?  W  W  P  -

                CCGAGTGCTGCANANGCTGTGCGANCGCGGCGCGAANAACGTGCTGGCCTTCGGCTTCGC
            301 ---------+---------+---------+---------+---------+---------+ 360
                GGCTCACGACGTNTNCGACACGCTNGCGCCGCGCTTNTTGCACGACCGGAAGCCGAAGCG

a                 P  S  A  A  ?  A  V  R  ?  R  R  E  ?  R  A  G  L  R  L  R  -
b                  R  V  L  ?  ?  L  C  ?  R  G  A  ?  N  V  L  A  F  G  F  A  -
c                   E  C  C  ?  ?  C  A  ?  A  A  R  ?  T  C  W  P  S  A  S  R  -

                GCTGCTGGACGGGGCCCGCGGGGGCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTA
            361 ---------+---------+---------+---------+---------+---------+ 420
                CGACGACCTGCCCCGGGCGCCCCCGGGGGGGCTCCGGAAGTGGTGGTCGCACGCGTCGAT

a                 A  A  G  R  G  P  R  G  P  P  R  G  L  H  H  Q  R  A  Q  L  -
b                  L  L  D  G  A  R  G  G  P  P  E  A  F  T  T  S  V  R  S  Y  -
c                   C  W  T  G  P  A  G  A  P  P  R  P  S  P  P  A  C  A  A  T  -

                CCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGCTGCG
            421 ---------+---------+---------+---------+---------+---------+ 480
                GGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTCGCCCCGCACCCCCGACGACGACGC

a                 P  A  Q  H  G  D  R  R  T  A  G  E  R  G  V  G  A  A  A  A  -
b                  L  P  N  T  V  T  D  A  L  R  G  S  G  A  W  G  L  L  L  R  -
c                   C  P  T  R  *  P  T  H  C  G  G  A  G  R  G  G  C  C  C  A  -
```

FIG. 72A

```
a      P R G R R R A G S P A G T L R ? ? C A G -
b       R V G D D V L V H L L A R C A ? F V L V -
c        A W A T T C W F T C W H A A R ? L C W W -

       GGNTCCCAGCTGCGCCTACCANGTGTGCGGGCCGCCGCTGTACCAGCTCGGCGCTGCNAC
541    ---------+---------+---------+---------+---------+---------+ 600
       CCNAGGGTCGACGCGGATGGTNCACACGCCCGGCGGCGACATGGTCGAGCCGCGACGNTG

a      G S Q L R L P ? V R A A A V P A R R C ? -
b       ? P S C A Y ? V C G P P L Y Q L G A A T -
c        ? P A A P T ? C A G R R C T S S A L ? L -

       TCAGGCCCGGCCCCCGCCACACGCTANTGGACCCGAANGCGTCTGGGATCCAACGGGCCT
601    ---------+---------+---------+---------+---------+---------+ 660
       AGTCCGGGCCGGGGGGCGGTGTGCGATNACCTGGGCTTNCGCAGACCCTAGGTTGCCCGGA

a      S G P A P A T R ? W T R ? R L G S N G P -
b       Q A R P P P H A ? G P E ? V W D P T G L -
c        R P G P R H T L ? D P ? A S G I Q R A W -

       GGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGGCTGCCAGCCCCGGGTGCGAGGAG
661    ---------+---------+---------+---------+---------+---------+ 720
       CCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCCGACGGTCGGGGCCCACGCTCCTC

a      G T I A S G R P G S P W A A S P G C E E -
b       E P * R Q G G R G P P G L P A P G A R R -
c        N H S V R E A G V P L G C Q P R V R G G -

       GCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCCAGGCGTGGCGCTGC
721    ---------+---------+---------+---------+---------+---------+ 780
       CGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGGGTTCTCCGGGTCCGCACCGCGACG

a      A R G Q C Q P K S A V A Q E A Q A W R C -
b       R G G S A S R S L P L P K R P R R G A A -
c        A G A V P A E V C R C P R G P G V A L P -

       CCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTCCTGGGCCCACCCGGGCAGGACGCC
781    ---------+---------+---------+---------+---------+---------+ 840
       GGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAGGACCCGGGTGGGCCCGTCCTGCGG

a      P * A G A D A R W A G V L G P P G Q D A -
b       P E P E R T P V G Q G S W A H P G R T P -
c        L S R S G R P L G R G P G P T R A G R L -

       TGGACCGAGTGACCGTGGTTTCTGTGTGGTGTCACCTGCCAGACCCGCCGAAGAAGCCAC
841    ---------+---------+---------+---------+---------+---------+ 900
       ACCTGGCTCACTGGCACCAAAGACACACCACAGTGGACGGTCTGGGCGGCTTCTTCGGTG

a      W T E * P W F L C G V T C Q T R R R S H -
b       G P S D R G F C V V S P A R P A E E A T -
c        D R V T V V S V W C H L P D P P K K P P -

       CTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCCGCCAGCA
901    ---------+---------+---------+---------+---------+---------+ 960
       GAGAAACCTCCCACGCGAGAGACCGTGCGCGGTGAGGGTGGGTAGGCACCCGGCGGTCGT

a      L F G G C A L W H A P L P P P I R G P P A -
b       S L E G A L S G T R H S H P S V G R Q H -
c        L W R V R S L A R A T P T H P W A A S T -

       CCACGCGGGCCCCCCATCCACATCGCGGCCACCACGTCCTGGGACACGCCTTGTCCCCCG
961    ---------+---------+---------+---------+---------+---------+ 1020
       GGTGCGCCCGGGGGGTAGGTGTAGCGCCGGTGGTGCAGGACCCTGTGCGGAACAGGGGGC
```

FIG. 72B

```
a       P  R  G  P  P  I  H  I  A  A  T  T  S  W  D  T  P  C  P  P  -
b        H  A  G  P  P  S  T  S  R  P  P  R  P  G  T  R  L  V  P  R  -
c         T  R  A  P  H  P  H  R  G  H  H  V  L  G  H  A  L  S  P  G  -

          GTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACAAGNACACTGCGNCCCTC
     1021 ---------+---------+---------+---------+---------+---------+ 1080
          CACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAGTCCGCTGTTCNTGTGACGCNGGGAG

a       V  Y  A  E  T  K  H  F  L  Y  S  S  G  D  K  ?  T  A  ?  L  -
b        C  T  P  R  P  S  T  S  S  T  P  Q  A  T  ?  T  L  R  P  S  -
c         V  R  R  D  Q  A  L  P  L  L  L  R  R  Q  ?  H  C  ?  P  P  -

          CTTCCTACTCAATATATCTGAGGCCCAGCCTGACTGGCGTTCGGGAGGTTCGTGGAGACA
     1081 ---------+---------+---------+---------+---------+---------+ 1140
          GAAGGATGAGTTATATAGACTCCGGGTCGGACTGACCGCAAGCCCTCCAAGCACCTCTGT

a       L  P  T  Q  Y  I  *  G  P  A  *  L  A  F  G  R  F  V  E  T  -
b        F  L  L  N  I  S  E  A  Q  P  D  W  R  S  G  G  S  W  R  ?  -
c         S  Y  S  I  Y  L  R  P  S  L  T  G  V  R  E  V  R  G  D  ?  -

          NTCTTTCTGGTTCCAGGCCTTGGATGCCAGGATTCCCCGCAGGTTGCCCCGCCTGCCCCA
     1141 ---------+---------+---------+---------+---------+---------+ 1200
          NAGAAAGACCAAGGTCCGGAACCTACGGTCCTAAGGGGCGTCCAACGGGGCGGACGGGGT

a       ?  F  L  V  P  G  L  G  C  Q  D  S  P  Q  V  A  P  P  A  P  -
b        S  F  W  F  Q  A  L  D  A  R  I  P  R  R  L  P  R  L  P  Q  -
c         L  S  G  S  R  P  W  M  P  G  F  P  A  G  C  P  A  C  P  S  -

          GCGNTACTGGCAAATGCGGCCCCTGTTTCTGGAGCTGCTTGGGAACCACGCGCAGTGCCC
     1201 ---------+---------+---------+---------+---------+---------+ 1260
          CGCNATGACCGTTTACGCCGGGGACAAAGACCTCGACGAACCCTTGGTGCGCGTCACGGG

a       A  ?  L  A  N  A  A  P  V  S  G  A  A  W  E  P  R  A  V  P  -
b        R  Y  W  Q  M  R  P  L  F  L  E  L  L  G  N  H  A  Q  C  P  -
c         ?  T  G  K  C  G  P  C  F  W  S  C  L  G  T  T  R  S  A  P  -

          CTACGGGGTGTTCCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTCACCCCAGCAGCCGG
     1261 ---------+---------+---------+---------+---------+---------+ 1320
          GATGCCCCACAAGGAGTTCTGCGTGACGGGCGACGCTCGACGCCAGTGGGGTCGTCGGCC

a       L  R  G  V  P  Q  D  A  L  P  A  A  S  C  G  H  P  S  S  R  -
b        Y  G  V  F  L  K  T  H  C  P  L  R  A  A  V  T  P  A  A  G  -
c         T  G  C  S  S  R  R  T  A  R  C  E  L  R  S  P  Q  Q  P  V  -

          TGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGAACACAG
     1321 ---------+---------+---------+---------+---------+---------+ 1380
          ACAGACACGGGCCCTCTTCGGGGTCCCGAGACACCGCCGGGGGCTCCTCCTCCTTGTGTC

a       C  L  C  P  G  E  A  P  G  L  C  G  G  P  R  G  G  G  T  Q  -
b        V  C  A  R  E  K  P  Q  G  S  V  A  A  P  E  E  E  E  H  R  -
c         S  V  P  G  R  S  P  R  A  L  W  R  P  P  R  R  R  N  T  D  -

          ACCCCCGTCGCCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCCTGGCAGGTGTACGGCT
     1381 ---------+---------+---------+---------+---------+---------+ 1440
          TGGGGGCAGCGGACCACGTCGACGAGGCGGTCGTGTCGTCGGGGACCGTCCACATGCCGA

a       T  P  V  A  W  C  S  C  S  A  S  T  A  A  P  G  R  C  T  A  -
b        P  P  S  P  G  A  A  A  P  P  A  Q  Q  P  L  A  G  V  R  L  -
c         P  R  R  L  V  Q  L  L  R  Q  H  S  S  P  W  Q  V  Y  G  F  -

          TCGTGCCGGGCCTGCCTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGGCACAACG
     1441 ---------+---------+---------+---------+---------+---------+ 1500
          AGCACGCCCGGACGGACGCGGCCGACCACGGGGGTCCGGAGACCCCGAGGTCCGTGTTGC
```

FIG. 72C

```
a       S C G P A C A G W C P Q A S G A P G T T -
b        R A G L P A P A G A P R P L G L Q A Q R -
c         V R A C L R R L V P P G L W G S R H N E -

        AACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCT
   1501 ---------+---------+---------+---------+---------+---------+ 1560
        TTGCGGCGAAGGAGTCCTTGTGGTTCTTCAAGTAGAGGGACCCCTTCGTACGGTTCGAGA

a       N A A S S G T P R S S S P W G S M P S S -
b        T P L P Q E H Q E V H L P G E A C Q A L -
c         R R F L R N T K K F I S L G K H A K L S -

        CGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGCTTGGCTGCGCAGGAGCC
   1561 ---------+---------+---------+---------+---------+---------+ 1620
        GCGACGTCCTCGACTGCACCTTCTACTCGCACGCCCTGACGCGAACCGACGCGTCCTCGG

a       R C R S * R G R * A C G T A L G C A G A -
b        A A G A D V E D E R A G L R L A A Q E P -
c         L Q E L T W K M S V R D C A W L R R S P -

        CAGGGGTTGGCTGTGTTCCGGCCGCAGAGCACCGTCTGCGTGAGGAGATCCTGGCCAAGT
   1621 ---------+---------+---------+---------+---------+---------+ 1680
        GTCCCCAACCGACACAAGGCCGGCGTCTCGTGGCAGACGCACTCCTCTAGGACCGGTTCA

a       Q G L A V F R P Q S T V C V R R S W P S -
b        R G W L C S G R R A P S A * G D P G Q V -
c         G V G C V P A A E H R L R E E I L A K F -

        TCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCA
   1681 ---------+---------+---------+---------+---------+---------+ 1740
        AGGACGTGACCGACTACTCACACATGCAGCAGCTCGACGAGTCCAGAAAGAAAATACAGT

a       S C T G * * V C T S S S C S G L S F M S -
b        P A L A D E C V R R R A A Q V F L L C H -
c         L H W L M S V Y V V E L L R S F F Y V T -

        CGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTGGAGCAAGT
   1741 ---------+---------+---------+---------+---------+---------+ 1800
        GCCTCTGGTGCAAAGTTTTCTTGTCCGAGAAAAAGATGGCCTTCTCACAGACCTCGTTCA

a       R R P R F K R T G S F S T G R V S G A S -
b        G D H V S K E Q A L F L P E E C L E Q V -
c         E T T F Q K N R L F F Y R K S V W S K L -

        TGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTCGGAAG
   1801 ---------+---------+---------+---------+---------+---------+ 1860
        ACGTTTCGTAACCTTAGTCTGTCGTGAACTTCTCCCACGTCGACGCCCTCGACAGCCTTC

a       C K A L E S D S T * R G C S C G S C R K -
b        A K H W N Q T A L E E G A A A G A V G S -
c         Q S I G I R Q H L K R V Q L R E L S E A -

        CAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCA
   1861 ---------+---------+---------+---------+---------+---------+ 1920
        GTCTCCAGTCCGTCGTAGCCCTTCGGTCCGGGCGGGACGACTGCAGGTCTGAGGCGAAGT

a       Q R S G S I G K P G P P C * R P D S A S -
b        R G Q A A S G S Q A R P A D V Q T P L H -
c         E V R Q H R E A R P A L L T S R L R F I -

        TCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAA
   1921 ---------+---------+---------+---------+---------+---------+ 1980
        AGGGGTTCGGACTGCCCGACGCCGGCTAACACTTGTACCTGATGCAGCACCCTCGGTCTT
```

FIG. 72D

```
a       S  P  S  L  T  G  C  G  R  L  *  T  W  T  T  S  W  E  P  E  -
b        P  Q  A  *  R  A  A  A  D  C  E  H  G  L  R  R  G  S  Q  N  -
c         P  K  P  D  G  L  R  P  I  V  N  M  D  Y  V  V  G  A  R  T  -

        CGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGCG
1981    ---------+---------+---------+---------+---------+---------+---------+ 2040
        GCAAGGCGTCTCTTTTCTCCCGGCTCGCAGAGTGGAGCTCCCACTTCCGTGACAAGTCGC

a       R  S  A  E  K  R  G  P  S  V  S  P  R  G  *  R  H  C  S  A  -
b        V  P  Q  R  K  E  G  R  A  S  H  L  E  G  E  G  T  V  Q  R  -
c         F  R  R  E  K  R  A  E  R  L  T  S  R  V  K  A  L  F  S  V  -

        TGCTCAACTACGAGCGGGCGCGGCGCCCCGGCCTCCTGGGCGCCTCTGTGCTGGGCCTGG
2041    ---------+---------+---------+---------+---------+---------+---------+ 2100
        ACGAGTTGATGCTCGCCCGCGCCGCGGGGCCGGAGGACCCGCGGAGACACGACCCGGACC

a       C  S  T  T  S  G  R  G  A  P  A  S  W  A  P  L  C  W  A  W  -
b        A  Q  L  R  A  G  A  A  P  R  P  P  G  R  L  C  A  G  P  G  -
c         L  N  Y  E  R  A  R  R  P  G  L  L  G  A  S  V  L  G  L  D  -

        ACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGACCCGCCGC
2101    ---------+---------+---------+---------+---------+---------+---------+ 2160
        TGCTATAGGTGTCCCGGACCGCGTGGAAGCACGACGCACACGCCCGGGTCCTGGGCGGCG

a       T  I  S  T  G  P  G  A  P  S  C  C  V  C  G  P  R  T  R  R  -
b        R  Y  P  Q  G  L  A  H  L  R  A  A  C  A  G  P  G  P  A  A  -
c         D  I  H  R  A  W  R  T  F  V  L  R  V  R  A  Q  D  P  P  P  -

        CTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCCAGGACA
2161    ---------+---------+---------+---------+---------+---------+---------+ 2220
        GACTCGACATGAAACAGTTCCACCTACACTGCCCGCGCATGCTGTGGTAGGGGGTCCTGT

a       L  S  C  T  L  S  R  W  M  *  R  A  R  T  T  P  S  P  R  T  -
b        *  A  V  L  C  Q  G  G  C  D  G  R  V  R  H  H  P  P  G  Q  -
c         E  L  Y  F  V  K  V  D  V  T  G  A  Y  D  T  I  P  Q  D  R  -

        GGCTCACGGAGGTCATCGCCAGCATCATCAAACCCCAGAACACGTACTGCGTGCGTCGGT
2221    ---------+---------+---------+---------+---------+---------+---------+ 2280
        CCGAGTGCCTCCAGTAGCGGTCGTAGTAGTTTGGGGTCTTGTGCATGACGCACGCAGCCA

a       G  S  R  R  S  S  P  A  S  S  N  P  R  T  R  T  A  C  V  G  -
b        A  H  G  G  H  R  Q  H  H  Q  T  P  E  H  V  L  R  A  S  V  -
c         L  T  E  V  I  A  S  I  I  K  P  Q  N  T  Y  C  V  R  R  Y  -

        ATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAGCCACGTCT
2281    ---------+---------+---------+---------+---------+---------+---------+ 2340
        TACGGCACCAGGTCTTCCGGCGGGTACCCGTGCAGGCGTTCCGGAAGTTCTCGGTGCAGA

a       M  P  W  S  R  R  P  P  M  G  T  S  A  R  P  S  R  A  T  S  -
b        C  R  G  P  E  G  R  P  W  A  R  P  Q  G  L  Q  E  P  R  L  -
c         A  V  V  Q  K  A  A  H  G  H  V  R  K  A  F  K  S  H  V  S  -

        CTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGANAACA
2341    ---------+---------+---------+---------+---------+---------+---------+ 2400
        GATGGAACTGTCTGGAGGTCGGCATGTACGCTGTCAAGCACCGAGTGGACGTCCTNTTGT

a       L  P  *  Q  T  S  S  R  T  C  D  S  S  W  L  T  C  R  ?  T  -
b        Y  L  D  R  P  P  A  V  H  A  T  V  R  G  S  P  A  G  ?  Q  -
c         T  L  T  D  L  Q  P  Y  M  R  Q  F  V  A  H  L  Q  ?  N  S  -

        GCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGCAGTG
2401    ---------+---------+---------+---------+---------+---------+---------+ 2460
        CGGGCGACTCCCTACGGCAGCAGTAGCTCGTCTCGAGGAGGGACTTACTCCGGTCGTCAC
```

FIG. 72E

```
a       A R * G M P S S S S R A P P * M R P A V -
b        P A E G C R R H R A E L L P E * G Q Q W -
c         P L R D A V V I E Q S S S L N E A S S G -

        GCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGT
   2461 ---------+---------+---------+---------+---------+---------+ 2520
        CGGAGAAGCTGCAGAAGGATGCGAAGTACACGGTGGTGCGGCACGCGTAGTCCCCGTTCA

a       A S S T S S Y A S C A T T P C A S G A S -
b        P L R R L P T L H V P P R R A H Q G Q V -
c         L F D V F L R F M C H H A V R I R G K S -

        CCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCC
   2521 ---------+---------+---------+---------+---------+---------+ 2580
        GGATGCAGGTCACGGTCCCCTAGGGCGTCCCGAGGTAGGAGAGGTGCGACGAGACGTCGG

a       P T S S A R G S R R A P S S P R C S A A -
b        L R P V P G D P A G L H P L H A A L Q P -
c         Y V Q C Q G I P Q G S I L S T L L C S L -

        TGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGACGGGCTGCTCC
   2581 ---------+---------+---------+---------+---------+---------+ 2640
        ACACGATGCCGCTGTACCTCTTGTTCGACAAACGCCCCTAAGCCGCCCTGCCCGACGAGG

a       C A T A T W R T S C L R G F G G T G C S -
b        V L R R H G E Q A V C G D S A G R A A P -
c         C Y G D M E N K L F A G I R R D G L L L -

        TGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAACCTTCC
   2641 ---------+---------+---------+---------+---------+---------+ 2700
        ACGCAAACCACCTACTAAAGAACAACCACTGTGGAGTGGAGTGGGTGCGCTTTTGGAAGG

a       C V W W M I S C W * H L T S P T R K P S -
b        A F G G * F L V G D T S P H P R E N L P -
c         R L V D D F L L V T P H L T H A K T F L -

        TCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACTTGCGGAAGACAG
   2701 ---------+---------+---------+---------+---------+---------+ 2760
        AGTCCTGGGACCAGGCTCCACAGGGACTCATACCGACGCACCACTTGAACGCCTTCTGTC

a       S G P W S E V S L S M A A W * T C G R Q -
b        Q D P G P R C P * V W L R G E L A E D S -
c         R T L V R G V P E Y G C V V N L R K T V -

        TGGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGATGCCGG
   2761 ---------+---------+---------+---------+---------+---------+ 2820
        ACCACTTGAAGGGACATCTTCTGCTCCGGGACCCACCGTGCCGAAAACAAGTCTACGGCC

a       W * T S L * K T R P W V A R L L F R C R -
b        G E L P C R R R G P G W H G F C S D A G -
c         V N F P V E D E A L G G T A F V Q M P A -

        CCCACGGCCTATTCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGA
   2821 ---------+---------+---------+---------+---------+---------+ 2880
        GGGTGCCGGATAAGGGGACCACGCCGGACGACGACCTATGGGCCTGGGACCTCCACGTCT

a       P T A Y S P G A A C C W I P G P W R C R -
b        P R P I P L V R P A A G Y P D P G G A E -
c         H G L F P W C G L L L D T R T L E V Q S -

        GCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCT
   2881 ---------+---------+---------+---------+---------+---------+ 2940
        CGCTGATGAGGTCGATACGGGCCTGGAGGTAGTCTCGGTCAGAGTGGAAGTTGGCGCCGA
```

FIG. 72F

```
a       A T T P A M P G P P S E P V S P S T A A -
b        R L L Q L C P D L H Q S Q S H L Q P R L -
c         D Y S S Y A R T S I R A S L T F N R G F -

        TCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGTGTCACA
2941    ---------+---------+---------+---------+---------+---------+ 3000
        AGTTCCGACCCTCCTTGTACGCAGCGTTTGAGAAACCCCAGAACGCCGACTTCACAGTGT

a       S R L G G T C V A N S L G S C G * S V T -
b        Q G W E E H A S Q T L W G L A A E V S Q -
c         K A G R N M R R K L F G V L R L K C H S -

        GCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTACAAGA
3001    ---------+---------+---------+---------+---------+---------+ 3060
        CGGACAAAGACCTAAACGTCCACTTGTCGGAGGTCTGCCACACGTGGTTGTAGATGTTCT

a       A C F W I C R * T A S R R C A P T S T R -
b        P V S G F A G E Q P P D G V H Q H L Q D -
c         L F L D L Q V N S L Q T V C T N I Y K I -

        TCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTCATCAGC
3061    ---------+---------+---------+---------+---------+---------+ 3120
        AGGAGGACGACGTCCGCATGTCCAAAGTGCGTACACACGACGTCGAGGGTAAAGTAGTCG

a       S S C C R R T G F T H V C C S S H F I S -
b        P P A A G V Q V S R M C A A A P I S S A -
c         L L L Q A Y R F H A C V L Q L P F H Q Q -

        AAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGCCTCCCTCTGCT
3121    ---------+---------+---------+---------+---------+---------+ 3180
        TTCAAACCTTCTTGGGGTGTAAAAAGGACGCGCAGTAGAGACTGTGCCGGAGGGAGACGA

a       K F G R T P H F S C A S S L T R P P S A -
b        S L E E P H I F P A R H L * H G L P L L -
c         V W K N P T F F L R V I S D T A S L C Y -

        ACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGCCAAGGGCGCCGCCGGCC
3181    ---------+---------+---------+---------+---------+---------+ 3240
        TGAGGTAGGACTTTCGGTTCTTGCGTCCCTACAGCGACCCCCGGTTCCCGCGGCGGCCGG

a       T P S * K P R T Q G C R W G P R A P P A -
b        L H P E S Q E R R D V A G G Q G R R R P -
c         S I L K A K N A G M S L G A K G A A G P -

        CTCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGCTGACTC
3241    ---------+---------+---------+---------+---------+---------+ 3300
        GAGACGGGAGGCTCCGGCACGTCACCGACACGGTGGTTCGTAAGGACGAGTTCGACTGAG

a       L C P P R P C S G C A T K H S C S S * L -
b        S A L R G R A V A V P P S I P A Q A D S -
c         L P S E A V Q W L C H Q A F L L K L T R -

        GACACCGTGTCACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGCAGCTGA
3301    ---------+---------+---------+---------+---------+---------+ 3360
        CTGTGGCACAGTGGATGCACGGTGAGGACCCCAGTGAGTCCTGTCGGGTCTGCGTCGACT

a       D T V S P T C H S W G H S G Q P R R S * -
b        T P C H L R A T P G V T Q D S P D A A E -
c         H R V T Y V P L L G S L R T A Q T Q L S -

        GTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGGCACTGC
3361    ---------+---------+---------+---------+---------+---------+ 3420
        CAGCCTTCGAGGGCCCCTGCTGCGACTGACGGGACCTCCGGCGTCGGTTGGGCCGTGACG
```

FIG. 72G

```
a       V  G  S  S  R  G  R  R  *  L  P  W  R  P  Q  P  T  R  H  C  -
b        S  E  A  P  G  D  D  A  D  C  P  G  G  R  S  Q  P  G  T  A  -
c         R  K  L  P  G  T  T  L  T  A  L  E  A  A  A  N  P  A  L  P  -

          CCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCGAGAGCA
3421      ---------+---------+---------+---------+---------+---------+ 3480
          GGAGTCTGAAGTTCTGGTAGGACCTGACTACCGGTGGGCGGGTGTCGGTCCGGCTCTCGT

a       P  Q  T  S  R  P  S  W  T  D  G  H  P  P  T  A  R  P  R  A  -
b        L  R  L  Q  D  H  P  G  L  M  A  T  R  P  Q  P  G  R  E  Q  -
c         S  D  F  K  T  I  L  D  *  W  P  P  A  H  S  Q  A  E  S  R  -

          GACACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCCACAC
3481      ---------+---------+---------+---------+---------+---------+ 3540
          CTGTGGTCGTCGGGACAGTGCGGCCCGAGATGCAGGGTCCCTCCCTCCCCGCCGGGTGTG

a       D  T  S  S  P  V  T  P  G  S  T  S  Q  G  G  R  G  G  P  H  -
b        T  P  A  A  L  S  R  R  A  L  R  P  R  E  G  G  A  A  H  T  -
c         H  Q  Q  P  C  H  A  G  L  Y  V  P  G  R  E  G  R  P  T  P  -

          CCAGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGT
3541      ---------+---------+---------+---------+---------+---------+ 3600
          GGTCCGGGCGTGGCGACCCTCAGACTCCGGACTCACTCACAAACCGGCTCCGGACGTACA

a       P  G  P  H  R  W  E  S  E  A  *  V  S  V  W  P  R  P  A  C  -
b        Q  A  R  T  A  G  S  L  R  P  E  *  V  F  G  R  G  L  H  V  -
c         R  P  A  P  L  G  V  *  G  L  S  E  C  L  A  E  A  C  M  S  -

          CCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTG
3601      ---------+---------+---------+---------+---------+---------+ 3660
          GGCCGACTTCCGACTCACAGGCCGACTCCGGACTCGCTCACAGGTCGGTTCCCGACTCAC

a       P  A  E  G  *  V  S  G  *  G  L  S  E  C  P  A  K  G  *  V  -
b        R  L  K  A  E  C  P  A  E  A  *  A  S  V  Q  P  R  A  E  C  -
c         G  *  R  L  S  V  R  L  R  P  E  R  V  S  S  Q  G  L  S  V  -

          TCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCCAGGGCC
3661      ---------+---------+---------+---------+---------+---------+ 3720
          AGGTCGTGTGGACGGCAGAAGTGAAGGGGTGTCCGACCGCGAGCCGAGGTGGGGTCCCGG

a       S  S  T  P  A  V  F  T  S  P  Q  A  G  A  R  L  H  P  R  A  -
b        P  A  H  L  P  S  S  L  P  H  R  L  A  L  G  S  T  P  G  P  -
c         Q  H  T  C  R  L  H  F  P  T  G  W  R  S  A  P  P  Q  G  Q  -

          AGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGA
3721      ---------+---------+---------+---------+---------+---------+ 3780
          TCGAAAAGGAGTGGTCCTCGGGCCGAAGGTGAGGGGTGTATCCTTATCAGGTAGGGGTCT

a       S  F  S  S  P  G  A  R  L  P  L  P  T  *  E  *  S  I  P  R  -
b        A  F  P  H  Q  E  P  G  F  H  S  P  H  R  N  S  P  S  P  D  -
c         L  F  L  T  R  S  P  A  S  T  P  H  I  G  I  V  H  P  Q  I  -

          TTCGCCATTGTTCACCCCTCGCCCTGCCCTCCTTTGCCTTCCACCCCCACCATCCAGGTG
3781      ---------+---------+---------+---------+---------+---------+ 3840
          AAGCGGTAACAAGTGGGGAGCGGGACGGGAGGAAACGGAAGGTGGGGGTGGTAGGTCCAC

a       F  A  I  V  H  P  S  P  C  P  P  L  P  S  T  P  T  I  Q  V  -
b        S  P  L  F  T  P  R  P  A  L  L  C  L  P  P  P  P  S  R  W  -
c         R  H  C  S  P  L  A  L  P  S  F  A  F  H  P  H  H  P  G  G  -

          GAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTG
3841      ---------+---------+---------+---------+---------+---------+ 3900
          CTCTGGGACTCTTCCTGGGACCCTCGAGACCCTTAAACCTCACTGGTTTCCACACGGGAC
```

FIG. 72H

```
a       E  T  L  R  R  T  L  G  A  L  G  I  W  S  D  Q  R  C  A  L  -
b        R  P  *  E  G  P  W  E  L  W  E  F  G  V  T  K  G  V  P  C  -
c         D  P  E  K  D  P  G  S  S  G  N  L  E  *  P  K  V  C  P  V  -

        TACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGT
3901    ---------+---------+---------+---------+---------+---------+ 3960
        ATGTGTCCGCTCCTGGGACGTGGACCTACCCCCAGGGACACCCAGTTTAACCCCCCTCCA

a       Y  T  G  E  D  P  A  P  G  W  G  S  L  W  V  K  L  G  G  G  -
b        T  Q  A  R  T  L  H  L  D  G  G  P  C  G  S  N  W  G  E  V  -
c         H  R  R  G  P  C  T  W  M  G  V  P  V  G  Q  I  G  G  R  C  -

        GCTGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTGAAAAAAAAAAAAAAAAA
3961    ---------+---------+---------+---------+---------+---------+ 4020
        CGACACCCTCATTTTATGACTTATATACTCAAAAAGTCAAAACTTTTTTTTTTTTTTTTT

a       A  V  G  V  K  Y  *  I  Y  E  F  F  S  F  E  K  K  K  K  K  -
b        L  W  E  *  N  T  E  Y  M  S  F  S  V  L  K  K  K  K  K  K  -
c         C  G  S  K  I  L  N  I  *  V  F  Q  F  *  K  K  K  K  K  K  -

        AAAAAAAAA
4021    --------- 4029
        TTTTTTTTT

a       K  K  K  -
b        K  K  -
c         K  K  -
```

FIG. 72I

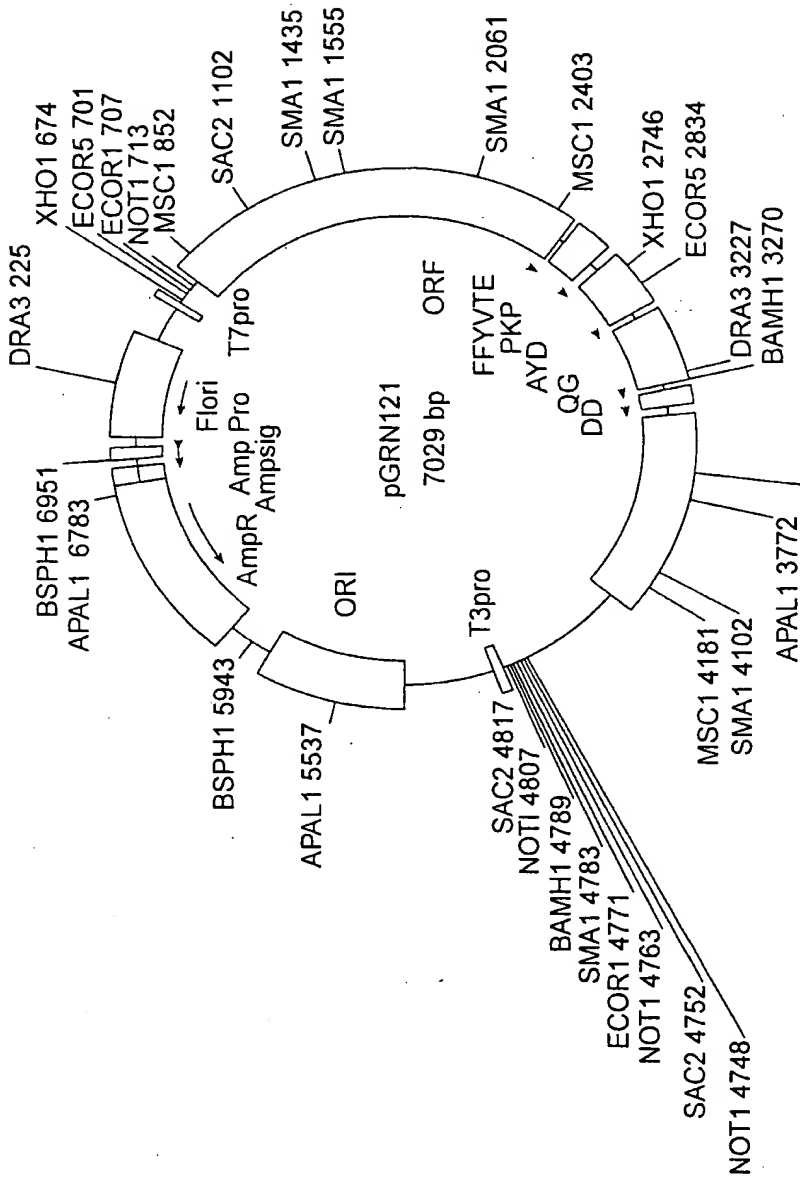


FIG. 73

```
                                                                      1
                                                                     met
GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG

                                        10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

                20                                              30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

                                        40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

                50                                              60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

                                        70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

                80                                              90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

                                       100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

               110                                             120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

                                       130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

               140                                             150
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

                                       160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

               170                                             180
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

                                       190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT
```

FIG. 74A

```
                200                                                 210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

                                    220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

                230                                             240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

                                    250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

                260                                             270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

                                    280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG
                290                                             300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

                                    310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

                320                                             330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

                                    340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

                350                                             360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

                                    370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

                380                                             390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

                                    400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

                410                                             420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA
```

FIG. 74B

```
                                    430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

            440                                         450
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

                                460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

            470                                         480
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

                                490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

            500                                         510
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

                                520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

            530                                         540
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

                                550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

            560                                         570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

                                580
arg leu phe phe tyr arg pro ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG CCG AGT GTC TGG AGC AAG TTG CAA AGC

            590                                         600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG
                                610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

            620                                         630
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

                                640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC
```

FIG. 74C

```
                650                                                 660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

                                    670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

            680                                                 690
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

                                    700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

            710                                                 720
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

                                    730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

            740                                                 750
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

                                    760
his gly his val arg lys ala phe lys ser his val ser thr leu
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG

            770                                                 780
thr asp leu gln pro tyr met arg gln phe val ala his leu gln
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG

                                    790
glu thr ser pro leu arg asp ala val val ile glu gln ser ser
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC

            800                                                 810
ser leu asn glu ala ser ser gly leu phe asp val phe leu arg
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC

                                    820
phe met cys his his ala val arg ile arg gly lys ser tyr val
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC

            830                                                 840
gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC

                                    850
cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG

            860                                                 870
ile arg arg asp gly leu leu leu arg leu val asp asp phe leu
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG
```

FIG. 74D

```
                                    880
leu val thr pro his leu thr his ala lys thr phe leu arg thr
TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

            890                                         900
leu val arg gly val pro glu tyr gly cys val val asn leu arg
CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG

                                    910
lys thr val val asn phe pro val glu asp glu ala leu gly gly
AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

            920                                         930
thr ala phe val gln met pro ala his gly leu phe pro trp cys
ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC
                                    940
gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
GGC CTG CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

            950                                         960
ser ser tyr ala arg thr ser ile arg ala ser val thr phe asn
TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT GTC ACC TTC AAC

                                    970
arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

            980                                         990
val leu arg leu lys cys his ser leu phe leu asp leu gln val
GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

                                    1000
asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

            1010                                        1020
leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

                                    1030
his gln gln val trp lys asn pro thr phe phe leu arg val ile
CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

            1040                                        1050
ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

                                    1060
ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GCC GGC CCT CTG CCC

            1070                                        1080
ser glu ala val gln trp leu cys his gln ala phe leu leu lys
TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

                                    1090
leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC
```

FIG. 74E

```
          1100                                                1110
arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

                                    1120
leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG GCA CTG CCC TCA GAC

            1130     1132
phe lys thr ile leu asp OP
TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCCACAGCCAGGCCGAGAGCAGA

CACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCCACACCC

AGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCC

GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC

CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAG

CTTTTCYTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT

CGCCATTGTTCACCCYTCGCCCTGCCYTCCTTTGCCTTCCACCCCCACCATCCAGGTGGA

GACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTA

CACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGC

TGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTGRAAAAAAAAAAAAAAAAAA

AAAAAAAAAA
```

FIG. 74F

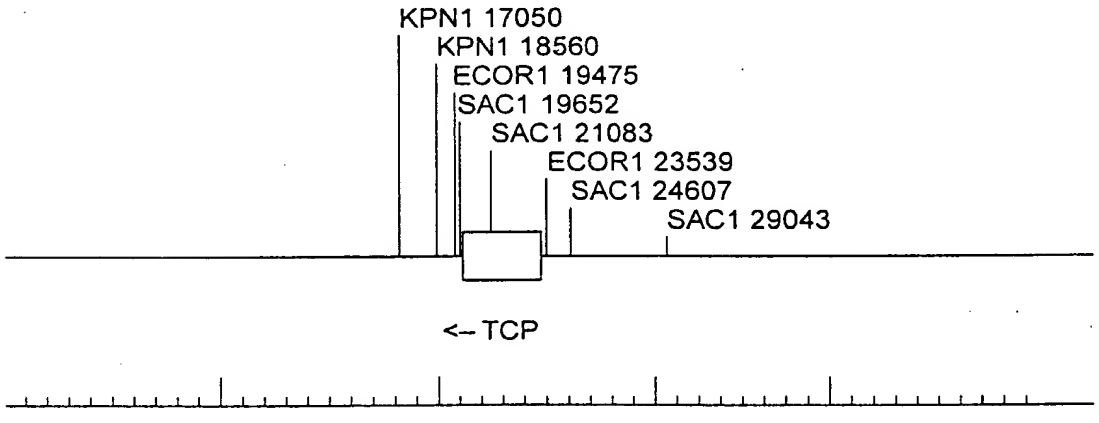


FIG. 75